From: Chan, Christina
Sent: Wednesday, May 10, 2006 5:46 PM
To: Graser, Jennifer; STIC-Biotech/ChemLib
Subject: RE: rush search

Please ~~just~~ Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----
From: Graser, Jennifer
Sent: Wednesday, May 10, 2006 5:46 PM
To: Chan, Christina
Subject: rush search
Importance: High

Hi Christina,

Will you please authorize the following rush search a special case which is due?

Thanks,
Jennifer

STIC:
Please search nucleotides 11357 to 12736 of SEQ ID NO: 35 from application 10/672787 in pending and commercial databases.
Thanks,
Jennifer Graser
Primary Examiner
Art Unit 1645
REM 3B09
2-0858

Na 96109

3C18

ME

OM nucleic - nucleic search, using sw model

```
                              GenCore version 5.1.8
                              Copyright (c) 1993 - 2006  Bioacceleration Ltd.

Run on:          May 18, 2006, 08:45:36 ; Search time 8124 Seconds
                                           (without alignments)
                                           10862.560 Million cell updates/sec

Title:           US-10-672-787-35_COPY_11357_12736
Perfect score:   1380
Sequence:        1 taaagcgtgaaccatgtaa..........acgacggcatattctttga 1380
Scoring table:   IDENTITY NUC
                 Gapop 10.0 , Gapext 1.0
Searched:        6366136 seqs, 3197371052S residues
Total number of hits satisfying chosen parameters:      11581394

Minimum DB seq length: 22
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       GenEmbl:*
                 1:  gb_env:*
                 2:  gb_pat:*
                 3:  gb_ph:*
                 4:  gb_pl:*
                 5:  gb_pr:*
                 6:  gb_ro:*
                 7:  gb_sts:*
                 8:  gb_sy:*
                 9:  gb_un:*
                 10: gb_vi:*
                 11: gb_ov:*
                 12: gb_htg:*
                 13: gb_in:*
                 14: gb_om:*
                 15: gb_ba:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	ID	Description
1	1380	100.0	96109	2	AR408756	AR408756 Sequence
2	1380	100.0	96109	2	AX067460	AX067460 Sequence
3	1370.4	99.3	1410	2	AX450009	AX450009 Sequence
4	266.8	19.3	110000	15	CP000082_25	Continuation (26 o
5	242	17.5	110000	15	CR543861_02	Continuation (3 of
6	229.6	16.6	110000	15	CP000082_24	Continuation (25 o
7	226.2	16.4	1416	2	AR318732	AR318732 Sequence
8	160.4	11.6	1314	2	AX318732	AX318732 Sequence
9	160.4	11.6	11498	2	AX932204	AX932204 Sequence
10	160.4	11.6	110000	2	AX191762	AX191762 Sequence
11	160.4	11.6	110000	2	BD426631_11	BD426631_12 Continuation (12 o
12	160.4	11.6	110000	2	BD426631_12	BD426631_11 Continuation (13 o
13	160.4	11.6	110000	2	AR274513_11	AR274513_11 Continuation (12 o
14	160.4	11.6	110000	2	AR274513_12	AR274513_12 Continuation (13 o
15	160.4	11.6	110000	2	AR632719_11	AR632719_11 Continuation (12 o
16	160.4	11.6	110000	2	AR632719_12	AR632719_12 Continuation (13 o
17	160.4	11.6	110000	15	L42023_11	L42023_11 Continuation (12 o
18	157.6	11.4	110000	15	CR628336_30	CR628336_30 Continuation (31 o
19	157.6	11.4	110000	15	CR628337_28	Continuation (29 o
20	154.8	11.2	110000	15	AE016828_01	Continuation (2 of
21	154.4	11.2	110000	2	CS222556	CS222556 Sequence
22	154.4	11.2	1311	2	CQ872956	CQ872956 Sequence
23	154.4	11.2	110000	15	CP000057_12	Continuation (13
24	154.4	11.2	349980	2	CQ873072	CQ873072 Sequence
25	154.2	11.2	110000	15	CP000155_59	Continuation (60 o
26	154.2	11.2	110000	15	CP000155_60	Continuation (61 o
27	152.8	11.1	110000	15	AE017354_29	Continuation (30 o
28	152.4	11.0	1314	2	AX932206	AX932206 Sequence
29	146.2	10.6	110000	15	BA000031_04	Continuation (5 of
30	142.2	10.3	12232	15	AE006048	AE006048 Pasteurel
31	139.6	10.1	1317	2	AX189058	AX189058 Sequence
32	139.6	10.1	2608	15	ECMUROY	ECMUROY Escherichia
33	139.6	10.1	28277	2	AX191720	AX191720 Sequence
34	139.6	10.1	28277	15	EC2MIN	EC2MIN E. coli 2 m
35	139.6	10.1	110000	15	U00096_00	U00096 Escherichia
36	139.6	10.1	110000	15	AE014075_01	Continuation (2 of
37	139.6	10.1	110000	15	AE014075_00	AE014075 Escherich
38	139.6	10.1	110000	15	AP009048_00	AP009048 Escherich
39	137.4	10.0	110000	15	AE017340_04	Continuation (5 of
40	137.4	10.0	110000	15	CP000034_02	Continuation (2 of
41	134.8	9.8	110000	15	AP009048_02	Continuation (2 of
42	134.8	9.8	110000	15	AE005174_01	Continuation (2 of
43	134.8	9.8	110000	15	AE005174_00	AE005174 Escherich
44	134.8	9.8	110000	15	BA000007_00	BA000007 Escherich
45	133.2	9.7	1629	15	X17609	X17609 Escherichia

ALIGNMENTS

```
RESULT 1
AR408756

LOCUS        AR408756        96109 bp    DNA    linear    PAT 18-DEC-2003
DEFINITION   Sequence 35 from patent US 6632636.
ACCESSION    AR408756
VERSION      AR408756.1    GI:40159151
KEYWORDS     .
SOURCE       Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 96109)
  AUTHORS    Lagace,R.E., Patterson,C., and Berg,K.L.
  TITLE      Nucleic acids encoding 3-ketoacyl-ACP reductase from Moraxella
             catarrhalis
  JOURNAL    Patent: US 6632636-A 35 14-OCT-2003;
             Elitra Pharmaceuticals Inc.; San Diego, CA
FEATURES             Location/Qualifiers
     source          1..96109
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match             100.0%;  Score 1380;  DB 2;  Length 96109;
  Best Local Similarity   100.0%;
  Matches 1380;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11357 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 11416

Qy         61 CATGCTAGCACAAGCACAGGTGACAAAAGCACAGCCCCTAGCGGTGCTTTGACTGGCTGGCCAA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11417 CATGCTAGCACAAGCACAGGTGACAAAAGCACAGCCCCTAGCGGTGCTTTGACTGGCTGGCCAA 11476

Qy        121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11477 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 11536

Qy        181 GAGTAAAATCTTGCTCAATTATCGGGGCCATCTTGACCGGATTAAATACACACTATCCACCCA 240
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11537 GAGTAAAATCTTGCTCAATTATCGGGGCCATCTTGACCGGATTAAATACACACTATCCACCA 11596
```

Bret8 22 amino acids

```
Qy          241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11597 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 11656

Qy          301 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11657 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 11716

Qy          361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11717 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 11776

Qy          421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11777 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 11836

Qy          481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11837 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 11896

Qy          541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCCATACC 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11897 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCCATACC 11956

Qy          601 ATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTTGTTGT 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11957 ATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTTGTTGT 12016

Qy          661 GATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12017 GATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 12076

Qy          721 GGCATCGTCTTGGTAGATGATGGCGGTTTGGCAATTTTCAAAAATACGCAATTTTTGTGC 780
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12077 GGCATCGTCTTGGTAGATGATGGCGGTTTGGCAATTTTCAAAAATACGCAATTTTTGTGC 12136

Qy          781 CAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTGT 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12137 CAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTGT 12196

Qy          841 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCAC 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12197 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCAC 12256

Qy          901 CACCAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTACCGCC 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12257 CACCAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTACCGCC 12316

Qy          961 GACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGCT 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12317 GACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGCT 12376

Qy         1021 TTTGGCATTTGAGCCTGTGATGGCGATGATTGGTGTGTTTGTTGCTTGAGTTAGATCTCG 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12377 TTTGGCATTTGAGCCTGTGATGGCGATGATTGGTGTGTTTGTTGCTTGAGTTAGATCTCG 12436

Qy         1081 TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTTT 1140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12437 TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTTT 12496

Qy         1141 GGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGGC 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12497 GGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGGC 12556

Qy         1201 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTTT 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12557 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTTT 12616

Qy         1261 ATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGAC 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12617 ATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGAC 12676

Qy         1321 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCATATTTTTTTGA 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        12677 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCATATTTTTTTGA 12736

RESULT 2
AX067460
LOCUS       AX067460               96109 bp    DNA     linear   PAT 24-JAN-2001
DEFINITION  Sequence 35 from Patent WO0078968.
ACCESSION   AX067460
VERSION     AX067460.1  GI:12545080
KEYWORDS    .
SOURCE      Moraxella catarrhalis
  ORGANISM  Moraxella catarrhalis
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Moraxellaceae; Moraxella.
REFERENCE   1
  AUTHORS   Lagace,R.E., Patterson,C. and Berg,K.L.
  TITLE     Nucleotide sequences of moraxella catarrhalis genome
  JOURNAL   Patent: WO 0078968-A 35 28-DEC-2000;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .96109
                     /organism="Moraxella catarrhalis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:480"
ORIGIN

  Query Match             100.0%;  Score 1380;  DB 2;  Length 96109;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1380;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy            1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11357 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 11416

Qy           61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11417 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 11476

Qy          121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11477 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 11536

Qy          181 GAGTAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACTATCCACCCA 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11537 GAGTAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACTATCCACCCA 11596

Qy          241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11597 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 11656

Qy          301 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11657 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 11716

Qy          361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11717 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 11776

Qy          421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11777 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 11836

Qy          481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11837 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 11896

Qy          541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCCATACC 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        11897 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCCATACC 11956

Qy          601 ATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTTGTTGT 660
```

```
RESULT 3
LOCUS       AR450009                1410 bp    DNA        linear   PAT 20-FEB-2004
DEFINITION  Sequence 673 from patent US 6673910.
ACCESSION   AR450009
VERSION     AR450009.1  GI:42679066
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1410)
  AUTHORS   Breton,G.L.
  TITLE     Nucleic acid and amino acid sequences relating to M. catarrhalis
            for diagnostics and therapeutics
  JOURNAL   Patent: US 6673910-A 673 06-JAN-2004;
            Genome Therapeutics Corporation; Waltham, MA
FEATURES             Location/Qualifiers
     source          1..1410
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

```
Query Match        99.3%;  Score 1370.4;  DB 2;  Length 1410;
Best Local Similarity   99.6%;  Pred. No. 0;  Matches 1374;  Conservative 0;  Mismatches 6;  Indels 0;  Gaps 0;
```

```
Qy      1 TAAAGCGTGAACCATTGTTAACAACATTATCACCTGCTCATTATATATTTTAAATTGATC    60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1392 TAAAGCGTGAACCATTGTTAACAACATTATCACCTGCTCATTATATATTTTAAATTGATC  1333

Qy     61 CATGTCTAGCACAACAAGCAGGTGACAAAAAGCACAGCCCCTTAGCCGTGCTTGACTGGCC   120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1332 CATGTCTAGCACAACAAGCAGGTGACAAAAAGCACAGCCCCTTAGCCGTGCTTGACTGGCC  1273

Qy    121 TACTGCCGGATCAAGTGTACCAACCTGATATATTATTTTAGAATGCCAAATGTGCATCACT   180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1272 TACTGCCGGATCAAGTGTACCAACCTGATATATTATTTTAGAATGCCAAATGTGCATCACT  1213

Qy    181 GAGTAAAATCTTGCTCAAATTATCGGGGGCCATCTTGACCGGATTAAAATACACACTATCA   240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1212 GAGTAAAATCTTGCTCAAATTATCGGGGGCCATCTTGACCGGATTAAAATACACACTATCA  1153

Qy    241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCCAACCCAAAAT       300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1152 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCCAACCCAAAAT     1093

Qy    301 CAATGCCAGTGATGATTATGCCATAAAACCGGTGCCAAGCCCATCAATCAATTGGTTGA    360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1092 CAATGCCAGTGATGATTATGCCATAAAACCGGTGCCAAGCCCATCAATCAATTGGTTGA   1033

Qy    361 GCCGATGTTGTACCTTTGAATCATTAAAAATATGCCTTGCCATCAATATCATCAATATA    420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1032 GCCGATGTTGTACCTTTGAATCATTAAAAATATGCCTTGCCATCAATATCATCAATATA    973

Qy    421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGGTATCAACATGCCATCCATGGG    480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    972 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGGTATCAACATGCCATCCATGGG    913

Qy    481 AAGATTGGGCAGCTGCTGGCCTAAGGCAAGTGCGAGATAAGGGCATTGAGTAGGTTATGC    540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    912 AAGATTGGGCAGCTGCTGGCCTAAGGCAAGTGCGAGATAAGGGCATTGAGTAGGTTATGC    853

Qy    541 TTTGATTTTAATTGATTGGTAAAAGTTGTTTTTTTTGATGATAAATGATAAATGCCATAC    600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    852 TTTGATTTTAATTGATTGGTAAAAGTTGTTTTTTTTGATGATAAATGATAAATGCCATAC    793

Qy    601 ATCACCCATCATCAAATAAATAAAAATCTGCCGTTTGGGTGGGCTAAGTAAGCTATTGTT    660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    792 ATCACCCATCATCAAATAAATAAAAATCTGCCGTTTGGGTGGGCTAAGTAAGCTATTGTT    733

Qy    661 GATAATGTGTGATTTGGGGCCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT    720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    732 GATAATGTGTGATTTGGGGCCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT    673

Qy    721 GGCATCGTCTTGGTAGATGAATGGGCGGGTTTGGCAATTTCAAATATACGCAATTTTTGT    780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    672 GGCATCGTCTTGGTAGATGAATGGGCGGGTTTGGCAATTTCAAATATACGCAATTTTTGT    613

Qy    781 CAAATAATCCACCATGCCATCGTGCCGATCAAGATGAGAGATTAAAATTGT          840
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    612 CAAATAATCCACCATGCCATCGTGCCGATCAAGATGAGAGATTAAAATTGT          553

Qy    841 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCAC   900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    552 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCAC   493

Qy    901 CACCAAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTACCGCC   960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    492 CACCAAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTACCGCC   433

Qy    961 GACACCAACAAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGCT  1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    432 GACACCAACAAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGCT   373
```

```
Qy    1021  TTTGGCATTTGAGCCTGTGATGGCGATGAGTTGTGTGTTGTTGCTTGAGTTAGATCTCG  1080
Db    372   TTTGGCATTTGAGCCTGTGATGGCGATGAGTTGTGTGTTGTTGCTTGAGTTAGATCTCG  313
Qy    1081  TGCTTTGAGTGTATCAAATGAACAAATTGAACATGCCTTGAGTTCGGATCTTT       1140
Db    312   TGCTTTGAGTGTATCAAATGAACAAATTGAACATGCCTTGAGTTCGGATCTTT       253
Qy    1141  GGCGGCGGACAATACTCGGCGGTTCTTGGGATTAATCGCGGGGCGTCGATCTGGC  1200
Db    252   GGGCGGCGGACAATACTCGGCGGTTCTTGGGATTAATCGCGGGGCGTCGATCTGGC  193
Qy    1201  TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGCAAGACAATGACGCGTTT  1260
Db    192   TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGCAAGACAATGACGCGTTT  133
Qy    1261  ATCTGCCAAGGTAGGATTAGGGATTACCATCGATAGTGCTAACTTTGTGACCGTGAATTGAC  1320
Db    132   ATCTGCCAAGGTAGGATTAGGGATTACCATCGATAGTGCTAACTTTGTGACCGTGAATTGAC  73
Qy    1321  CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCCAATATTTTTTTGA  1380
Db    72    CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCCAATATTTTTTTGA  13

RESULT 4
CP000082_25
WPCOMMENT
Sequence split into 27 fragments   LOCUS CP000082 Accession CP000082

Fragment Name    Begin      End
CP000082_00      1          110000
CP000082_01      100001     210000
CP000082_02      200001     310000
CP000082_03      300001     410000
CP000082_04      400001     510000
CP000082_05      500001     610000
CP000082_06      600001     710000
CP000082_07      700001     810000
CP000082_08      800001     910000
CP000082_09      900001     1010000
CP000082_10      1000001    1110000
CP000082_11      1100001    1210000
CP000082_12      1200001    1310000
CP000082_13      1300001    1410000
CP000082_14      1400001    1510000
CP000082_15      1500001    1610000
CP000082_16      1600001    1710000
CP000082_17      1700001    1810000
CP000082_18      1800001    1910000
CP000082_19      1900001    2010000
CP000082_20      2000001    2110000
CP000082_21      2100001    2210000
CP000082_22      2200001    2310000
CP000082_23      2300001    2410000
CP000082_24      2400001    2510000
CP000082_25      2500001    2610000
CP000082_26      2600001    2650701

Continuation (26 of 27) of CP000082 from base 2500001 (CP000082 Psychrobacter arcticus 2

Query Match       19.3%;    Score 266.8;  DB 15;   Length 110000;
Best Local Similarity   53.5%;    Pred. No. 2.1e-75;
Matches 745;   Conservative   0;   Mismatches 572;   Indels 75;   Gaps 6;

Qy    8     TGAACCATGTTAACAAACTTATCACCTCGCTCATTATTATATTTTTAAATTGATCCATGCTA  67
Db    8845  TTAACGAGCTGGCTAAAGCGCTCGCGCCCGCGGACAAACCCATTAAACTGATCGAGACTT  8904
Qy    68    GCACAAAGCAGGGTGACAAAAGCACAGCCCTTAGCGGTGCTTTGACTGGCGTCAATACTGCC  127
Db    8905  GCACAAAGCAGGGTGACAATAATACCCCGTGTTCGCAGAAAAAGTTTGCCGTTACTTGC  8964
Qy    128   CGA----------                                              128
Db    8965  TCAATCATCGAAAAAGCATTTTCTAAGGTCTGACATTGATGAGCAGGCTAACCTCATCGCTG  9024
Db    9025  AGACCGCGCTGCGGGTAAAAATGCCTTCAAATCGGGGGCATCTTGACCGATTAAAATACACA  9084
Qy    170   GCATCACCCTTTGAGTAAATCTTGCTCAATTATCGGGGGCATCTTGACCGATTAAAATACACA  229

Qy    230   CTATCCACCCACTGCATAACATATCTGACAGGCTCACTAAAATCTGTT       289
Db    9085  TGACTCACATATTGATTGAAAAAGGCGAGAGTTCACCGAACTGCTGG       9144
Qy    290   CCACCCAAAAATCAAATGCCA----------GTGATGATATGCCCAATAAAAACCGTCGCA  343
Db    9145  CCGCCCAAATATTAATAACTTACCGTCTTTGGGTGCATAGACCCGCCAAAACCTTCA  9204
Qy    344   ATTGCTGCAAATGGTTGAGCCGAAGTTGTACCTTTGAATCATTAATTATAAATGCCTTGCTA  403
Db    9205  ACAGCCGCCCAATGGTTGAGCCAGTAGGCCTTTAGAATCATTAATGTGCCAATTATCA  9264
Qy    404   TCAAATCATCAATATATATTCACAGCGGATGCGGGGTACCCCTTTGAAGTTCTTAAGGGTTA  463
Db    9265  GCAATTGTGGCGGACATATTGGCCGGATGCTTCAAGACCCCTTTGAAGTTCTTAAGGGGCTA  9324
Qy    464   AGCAATGGCATCCATGGGAACGACAGCAAGCTTCGCGCCAATGACTGCTTGCTTGGGCATTA  523
Db    9325  AGCAATCTCCAAGCGGGGCAAAACCAAGGCTCGCCCAATGGCAATGCTGCTTGGGCATTA  9384

Qy    524   AGTAGGTTATGCCTTGCCTTTGAATTTTTAATTGATGGATGGTAAAAGTTTGTTTTTTGA  583
Db    9385  AGTAAGTTATGACGACACTTTTATCTTTAGTTTATCGCTGATATGAGACGCTCTGTACCA  9444
Qy    584   TGATATAAATGCCATACCACCATCACAATAAAATCTCCGTTTTGGGGCTA         643
Db    9445  CGTGCCAAATAAGTCTGCCCCTGCGTCATCCGTTAATA------AGCCC---TGGG  9485

Qy    644   AGTAAGCTATTTGTTGTGATAATGGTGTGATTTGGGGGGCAACGCTTGTCAGTGTGTCAAG  703
Db    9486  ATAGTGCCCCTTGTTGGGACCATGAATACCAGTGTCAGTCGGGGCCAAAAGTTATCTCGACAC  9545
Qy    704   CATTGCTTGGCTAAAGTTGGCATCGTCTTTGGGTTGTAGATGGATCGAATGCAAA  763
Db    9546  CAATGGTCGGGTCAA--CGCATCTTCACGATTAAAGTTGGCATCTTCGATTGGGAAG  9603
Qy    764   ATACCGCAAATTAATAATCCACCATGCCATTCGTCGCCGATCAAGATGGATCAGCG  823
Db    9604  ATACGAATGGTTTGGCTTGGGTGGATAGCCGAGCAATATCACCATGACACCGTTTTGGGGCTA  9663
Qy    824   GAGAGATTTAAAAATTGTTTGCCCCTTGAGCTGCCAAAATTTGAGATGCTGAAA  883
Db    9664  GACATATTGGAGGAACTGTTGACAATCGGGTTGGTCACGTTCTTAACTGAAAG  9723
Qy    884   CTTGACAGCTGCAAGCACCACCCAAAATCCATATTTTGGATGTTGTTAATTCAAGTGCAGGC  943
Db    9724  CTAGAATAATTCTTATAAACTGCTAGCGTCCATTTTATCATTACTTATCGCCGCTGGC  9783
Qy    944   GTGCCAAATATTACCGCGACCACCAATACATGCCATGTTTTGCCATCTTACT  1003
Db    9784  ACGCCAAATATTACCGCCCTAGCCAACATTGATGCCGCCATTTGTCCCAACC  9843
Qy    1004  AATGTCGTACAGTGCTTTTTGGCATTTGACCCTGTCATGGGGATGATTGGTGTGTTTGTT  1063
Db    9844  ACGCTGGTCACAGTGCTCTTTTTGCATTTGAGCCAGTAACAATTCGGACCGTACAA  9903
Qy    1064  GCTTGAGTTAGATCTCGTCTGCTTGATTGTCAAATGAACAATGAACATCGCGATCACA  1123
Db    9904  GCT------------TCACAAAAATAGCTGATAATGCTGACCACT  9936
Qy    1124  GGAATGCCTTGAGCTTTTGGGGCGGCGACAAATACTCGGCGGTTAATGCCGGGGTG  1183
Db    9937  GGAATATTTGCCGCACGCGCGCGCCATCAATTATCTCGAGAAGGATAACGGACTG  9996
Qy    1184  ATGATGATCTGATCGGGCTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGAC---AA  1240
Db    9997  ATTATAAATCCGTGCCGCCTTGTTGTGTTAATAACTTCGATCGACGCACAAACTGTCTGATA  10056
```

```
Qy   1241 TGAACGCCTGTGTGGGGCAGTTTATTCTGCCCAAGGTAGGATTGGGATTACCATCGATGATGCTA 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10057 TCAATAGTAGCAGGCAGTTTATTCTGCTAATGCGGGATTCTTCCTGATTGTGTCAGTGACGGGCA 10116

Qy   1301 ACTTTGTTGACCGTGATTGACCAAAAAATTCACACGCTGACAGTCCAGAGCTGCCCAAACCG 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  10117 ACTTGGTTAGCCCTGCTCAGCCAAATAAATCCCGACCGGATTAAATCCCCCAAACCG 10176

Qy   1361 ACGACGGGATAT 1372
          ||||||||||||
Db  10177 ACCACGACTTGT 10188
```

RESULT 5

CR543861_02/c

WPCOMMENT

Sequence split into 36 fragments   LOCUS CR543861 Accession CR543861

Fragment Name	Begin	End
CR543861_00	1	110000
CR543861_01	100001	210000
CR543861_02	200001	310000
CR543861_03	300001	410000
CR543861_04	400001	510000
CR543861_05	500001	610000
CR543861_06	600001	710000
CR543861_07	700001	810000
CR543861_08	800001	910000
CR543861_09	900001	1010000
CR543861_10	1000001	1110000
CR543861_11	1100001	1210000
CR543861_12	1200001	1310000
CR543861_13	1300001	1410000
CR543861_14	1400001	1510000
CR543861_15	1500001	1610000
CR543861_16	1600001	1710000
CR543861_17	1700001	1810000
CR543861_18	1800001	1910000
CR543861_19	1900001	2010000
CR543861_20	2000001	2110000
CR543861_21	2100001	2210000
CR543861_22	2200001	2310000
CR543861_23	2300001	2410000
CR543861_24	2400001	2510000
CR543861_25	2500001	2610000
CR543861_26	2600001	2710000
CR543861_27	2700001	2810000
CR543861_28	2800001	2910000
CR543861_29	2900001	3010000
CR543861_30	3000001	3110000
CR543861_31	3100001	3210000
CR543861_32	3200001	3310000
CR543861_33	3300001	3410000
CR543861_34	3400001	3510000
CR543861_35	3500001	3599621

Continuation (3 of 36) of CR543861 from base 200001 (CR543861 Acinetobacter sp. ADP1 com

```
Query Match        17.5%;  Score 242;  DB 15;  Length 110000;
Best Local Similarity 52.9%;  Pred. No. 3.3e-67;
Matches 726;  Conservative 0;  Mismatches 580;  Indels 66;  Gaps 7;
```

```
Qy      5 GCCTGAACCATGTTAACAAACTTATCACCTCGCTCATTATATAATTTTTAAATTGATCCATG 64
          |||  |||||||||||      |||   ||| |||  ||||||||||||||||||||||
Db  71434 GCATCCACACAGCCAACAAAAATGATGTCCCGGATCCTGGATAAACCTTTAAACAATATCAAA 71375
```

```
Qy     65 CTAGCACAAGAGCAGGTGACAAAAGCACAGCCCTAGCGGGTGCTTGACTGGCTGCCAATCT 124
          ||||||| ||||| ||| ||||| |||||||||||||||||||||||||||||||||||
Db  71374 CTTGCACATGCAGGTGACAATAACACGAGTCGTTCTTCAGGCTGT----GCATACTGTTGT 71321
```

```
Qy    125 GCCCGATCAAGTGTTACCAAACCTGATATATTTTTAGATGCGATCACTTTTGAGT 184
          |||||||||||| ||||||||      |||  |||||||||||||||||||||||
Db  71320 GCCTGTGTCACCCGCATCTTTTAAAGTGGATGCATGCACCAACTGAGTTGTTTCTGT---C 71264
```

```
Qy   1259 TTATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTTGACCGTGATTG 1318
```

```
Qy    185 AAAATCTTGCTCAAATTATCGTGGGGCATCTTGACCGATTAAATACACTATCCACCCACTGC 244
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| || ||
Db  71263 AACGGCCTGCTCAATCCTGCTGAGCAATCTTCACCGATCAGAATCGCAGCCCTTCACATATTC 71204

Qy    245 ATAACAGAATATCGTGACCGTGACAGTCCAGAGCTCACAAAAATCGGGGATTCCACCCCAAAATC--- 301
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  71203 TCAACAGACTGGCGGCAGCAAGCTAAAAATCCTGGGCCTTTACCCTGCCGCCACCAAACCG 71144

Qy    302 ---AATGCCAGTGATGATATGCCATAAACGCGTTGCCCAAGCGCCATCAAATTGCTGCAATGGTT 358
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  71143 ATGAACTTTGCCTTGTTGTTGGGGTTCAATTGCAGCACCAAAGACCCATCAGCCAAGCGTT 71084

Qy    359 GAGCCGGATGTTTTGTACCTTTGAATCATTTCTAAAATATTTTACA 418
          |||||||||||||||||||||||||||||||||||||||||||
Db  71083 GCCCCCACATTGGTCCCTTTCGAATCATTATAATAACGTAACATTGACTGTTTTTACA 71024

Qy    419 TATTCACAGCGGATGCGGGTAGGCCCTTTGAAAGTCTTAAGGGTATCAAGCATACGACA 478
          |||||||||||||||||||||||||||||||||||||||||||
Db  71023 TATTCGGCAGGGATGCTCAAGCGCCCTTTAAACTGCTTCAAGCTGTTGCAAGGTATCGACA 70964

Qy    479 GGAAGATTGGCAGCGTGGCCTAAGGCAAGTGCAGATAAAGGCATTGAGTAGGTTATGCTTG 538
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  70963 GGTAAACCTATGCGTTCACCCCAATGCCAGACGAACAAGGCCAGTGCATTGGCAATAATTATGTGTTG 70904

Qy    539 CCTTTGATTTTTAAATTGATGGATTCGGTTGTAGAGATGGTTAAAAGTTTTGTTTTTTTGT 598
          |||||||||||||||||||||||||||||||||||||||||||
Db  70903 CCCTGAATATACAGCTCAGAACTCTTAATCAGGCGGGGTACGGCGTCAGTGCCACATG 70844

Qy    599 CCATCACCATCATCAAATAAATAAAATCTGCCGTTTGGGGGTTCCTAAGGTAAGCTATTGTT 658
          ||||||||||||||||||||||||||||||||||||||||||||
Db  70843 GTTCCATCATCTTCACGTAAAAACCGCCCATAATTGCTT------------AATGTCGG 70801

Qy    659 GTGATAATGGTGTTGATTGGGGCAACGCTTGTCAGTGTTGGTGGTTCAAGCATTGCTTGCTAA 718
          ||||||||||||||||||||||||||||||||||||||||||
Db  70800 GTGCATTGAGTCCAAAATCGCTTCACCCAATGCCAGAACGAAGCCAGTGCCATTGGCAATATCG 70741

Qy    719 TTGGCATCGTCTTTGGTTGTAGATGATGGTGCGGTTTGTTGTTTTTTTGTTCAAAAAATTAC 778
          |||||||||||||||||||||||||||||||||||||||
Db  70740 T----GTCATCACGGTTGTGTAAAACACCTTTTTACCCTGAAATTTACCCTGAAATTTACCG 70685

Qy    779 GCCAAATAATCCACCATGCCAATCGTGCCGATCAGATGATCAGCGGGAGGATTTAAAATT 838
          |||||||||||||||||||||||||||||||||||||||
Db  70684 TGATGATAGCCCAGATGTCTCCATGACGAATCTCACTCATATTTAATATG 70625

Qy    839 GTTGCCCCTTGAGCTGCCAAATTTGAGATAATGCTCAAGCTGAAAAACTTGACAGCTCAAGC 898
          |||||||||||||||||||||||||||||||||||||||
Db  70624 ACTGCCACCTCAGCATTCAGGTGCGAATGGTGTTCCAGTTGAAAAACTTCTAGT 70565

Qy    899 ACCACAAAATCCAATATTTTGGATGTTAATAATTCAAGTGCAGGGGGTGCCAATATTACCG 958
          |||||||||||||||||||||||||||||||||||||||
Db  70564 ACCACGAACTC---AGGCTCGTCTTTTAACAAAATCTAATCCTGGACGCCCTAAATTACCT 70508

Qy    959 CCGACACCAACAAATCCATGCCATGTTTTCCCATCCGCCTACTAAGTCGTAACAGTG 1018
          |||||||||||||||||||||||||||||||||||||||
Db  70507 CCTACTGCCACGCGTTACCTGCATCTTTTGCCATCTTTGACCAAACAAAGTGGGTTACGGTA 70448

Qy   1019 CTTTTGGCATTTGAGCGTTGCATGGCGATGATTGGTGTGTTGTTGCTTGAGTTGAGATCT 1078
          |||||||||||||||||||||||||||||||||||||||
Db  70447 CTTTTGGCATTTGAACCCGGTAAATCGGCAACAAATCGGAACGTCGGCACGCGGGT------ 70394

Qy   1079 CGTGCTTTGAGTGTATCAAATGAACATCGGTCGATCATCACAGGGAAGCCTTGTGCA 1138
          |||||||||||||||||||||||||||||||||||||||
Db  70393 ------------AAGACTTGAATATCTCCAAATGAACTGGGAATGCCTTGTGCA 70355

Qy   1139 TTGGGGGGGACAAATACTCGGGCGGTTCTTGGGTTAATTGTGTCCAGTCCTGATGGATTCTG 1198
          |||||||||||||||||||||||||||||||||||||||
Db  70354 ATCGGCTTGCTGTGGGATTTCGGGTAATTGTTGTGGCTTAAAAATAAATTTTCTTCG 70295

Qy   1199 GCTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAATGACAAATGAACGCCTTGTGTGAGT 1258
          |||||||||||||||||||||||||||||||||||||||
Db  70294 GCTTTGCAGGAGTAGTTGCAGTTATCTAACTGACCAAAAAACTGGTTTGCACATCTTTGAGGAATC 70235
```

```
Db    70234 TGATCATGCCCAGGTGGAGGTCTCTCGAGGAATCGTTGGACGAGCAGCAACGGCGGTAGCCTTGTTCA
                ||||                  |||||             |||   |||
Qy     1319 ACCAAAAAATTCACAGTCGGAGCTGCCAAACCGACGACGGCAT 1370

Db    70174 TGTAAAAAAATTAACGGCTGAAACCCTGAAGATTCCAAGACCCGCAACGACTT 70123


RESULT 6
WPCOMMENT
CP000082_24

Sequence split into 27 fragments        LOCUS  CP000082  Accession CP000082

Fragment Name    Begin        End
CP000082_00      1            110000
CP000082_01      100001       210000
CP000082_02      200001       310000
CP000082_03      300001       410000
CP000082_04      400001       510000
CP000082_05      500001       610000
CP000082_06      600001       710000
CP000082_07      700001       810000
CP000082_08      800001       910000
CP000082_09      900001       1010000
CP000082_10      1000001      1110000
CP000082_11      1100001      1210000
CP000082_12      1200001      1310000
CP000082_13      1300001      1410000
CP000082_14      1400001      1510000
CP000082_15      1500001      1610000
CP000082_16      1600001      1710000
CP000082_17      1700001      1810000
CP000082_18      1800001      1910000
CP000082_19      1900001      2010000
CP000082_20      2000001      2110000
CP000082_21      2100001      2210000
CP000082_22      2200001      2310000
CP000082_23      2300001      2410000
CP000082_24      2400001      2510000
CP000082_25      2500001      2610000
CP000082_26      2600001      2650701

Continuation (25 of 27) of CP000082 from base 2400001 (CP000082 Psychrobacter arcticus 2

Query Match    16.6%;    Score 229.6;    DB 15;    Length 110000;
Best Local Similarity    54.3%;    Pred. No. 4.1e-63;
Matches 581;    Conservative 0;    Mismatches 444;    Indels 45;    Gaps 4;
```

```
Qy       8 TGAACCATGTTAACAAAACTTATCACCTCGCTCATTATATATTTTAAATTGATCCATGCTA 67
             |||||||   ||| |||||||           ||||||||    ||||||||||||||
Db  108845 TTAACGAGTGTGGCTAAAGCGCTCGCCCCGCCGACAAACCCATTAAACTGATCGAAGCTT 108904

Qy      68 GCACAAGCAGGTGACAAAAGCACAGCCCTAGCGCGTGCTTTGACTGGCTGCCAATACTGCC 127
            |||||||||||||||||                        ||||            ||
Db  108905 GCACAAGCAGGTGACAATAATACCCGCTTGCCGACAAAACTACTTGCGCGTTACTTGC 108964

Qy     128 CGA------------------TCAAGTGGTACCAACCTGATATATATTTAGATGCCAAATGT 169
            |||                  |||
Db  108965 TCAATCATCGAAAAAGCATTTCTTAAGGTCTGACACATTGATGCAGCCTAACCTCATGCCTG 109024

Qy     170 GCATCACCTTTGAGTAAATCTTGCTCAATTATCGGGGCACTCGACGGATTAAATAACA 229
            |||          ||| |  |||||||||||         |||
Db  109025 AGACCCGCTGCGCGTAAATGCTGTTCAATCGTCCTCGCCAATAAAAGAGCACT 109084

Qy     230 CTATCCACCCACTCGCATACACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCCAAGC 289
            |||                |||           |||           |||
Db  109085 TGACTCACATATTGATTGAAAAAGGCAGAGAGTTCACCGGAACTGCTGAGCTTTACCTTGA 109144

Qy     290 CCACCCAAAATCAATGCCA------GTGATGATATGCGCAATAACCGTGCCAAGCCCATCA 343
            |||                      |||           |||           |||
Db  109145 CCGCCCAATATATTAATTAATAACTTACCGTCTTTGGGTGCATAGACACGGCGCCAAAACCTTCA 109204

Qy     344 ATTGCTGCAATGGTTGAGCCGATGTTTGTTGTACCTTTGTGAATCATTAAAAATATGCCTTGCTA 403
            ||||     |||||||||||||     |||||     |||||||||||||     ||||
Db  109205 ACAGCCGCCCATGGTTGAGCCAATATTGGTGCCTTTAGAATCATTAAAGTAATCAAATGCCA 109264
```

```
RESULT 7
AR318732/c
LOCUS        AR318732      1416 bp    DNA    linear    PAT 17-AUG-2003
DEFINITION   Sequence 1282 from patent US 6562958.
ACCESSION    AR318732
VERSION      AR318732.1  GI:33699835
KEYWORDS     .
SOURCE       Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 1416)
  AUTHORS    Breton,G. and Bush,D.
  TITLE      Nucleic acid and amino acid sequences relating to Acinetobacter
             baumannii for diagnostics and therapeutics
  JOURNAL    Patent: US 6562958-A 1282 13-MAY-2003;
             Genome Therapeutics Corporation; Waltham, MA
FEATURES             Location/Qualifiers
     source          1..1416
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match    16.4%;    Score 226.2;    DB 2;    Length 1416;
Best Local Similarity    52.1%;    Pred. No. 2.1e-62;
Matches 718;    Conservative 0;    Mismatches 593;    Indels 66;    Gaps 7;
```

```
Qy     404 TCAATATCATCAATATATATTCACAGCCGGATGGGGTAGCGCCCTTTGAAAGTTCTTAAGGGTGTATCA 463
             ||||  |||||    |||||||   ||||||| |||||      |||||      |||||
Db  109265 GCAAATTGTGGCGGACGCATTATTGGCCAGCCGGATGTGCCAAAACTGCCAAGTGCTA 109324

Qy     464 AGCATGGCATCATCGGGGAAGAGATTGGCAGCGGCTGTGCCTTAAGGCAAGTGCAGATAAGGCATTG 523
            |||               |||||             |||                  |||
Db  109325 AGCATACTCTCAAGCGGCAAAACCAGCAAGCTCGCCCAAATGCCAATGCTGCTTGCTTGGGCATTA 109384

Qy     524 AGTAGGTTATGCTTGCCTTTGATTTTTTAATTGATGAGTGTAAAAGTTTGTTTTTTTGA 583
            |||                                        |||
Db  109385 AGTAAGTTATGACGAGCGACTTTTATCTTTAGTTATTCGTGATGATATGAGAGACGGCTCTGTACCA 109444

Qy     584 TGATATAAATGCCATACCATCACCATCATCAAATAAATAAAATCTGCCGTTGTGGTGCTA 643
            |||                                                    |||
Db  109445 CGTGCCAAATAAGTCTGCCCTTGCCTGCATCCGTAATA------AGCCC 109485

Qy     644 AGTAAGCTATTTGTTGTGATAATGGTGATTGTGGGGCAACGGCTTGTCAGTGGTCAAG 703
            |||                                       |||
Db  109486 ATAGTGCCCCTTGTTAGGGACGCATGAATAACCAGTGCTCAGTGCGCAAAACTATCTGACAC 109545

Qy     704 CATTGCTTGGCTAAAATTGGCATCGTCTTGGTAGATGATGGGGTTGCAATTTCAAAA 763
            |||                                   |||
Db  109546 CAATGGTCGGGGTCAA--CGCATCTTCACGATTAATCACTACCGATTGCGAAGCTT 109603

Qy     764 ATACGCAATTTTTTGTGCCAAATAATCACCATCGTGCCGATCAAGATGATCAGCG 823
            |||                                                 |||
Db  109604 ATACGACGTTTGGCCTGGTGATAGCCGAGCAATATCACCATGAACGGTCTAAAATGATCAGGC 109663

Qy     824 GAGAGAGATTTAAAAATTGTTGCCCGTGCCAAAATTTGAGAGATAATGCTCAAAGCTGAAAA 883
            |||                                                       |||
Db  109664 GACATATTGAGGACGTTGGCGACCTTGCGCCCCTAGCTTGGGTCACGGTTTCTTAACCTGAAAG 109723

Qy     884 CTTGACAGCTCAAGCACCACCAAAATCCATATTTGGATGCGTTAATTAATTCAAGTGCGAGGC 943
            |||                                                        |||
Db  109724 CTAGATAATTCTATTAACTGCTAGCCTCCATTTTATCAAAAAGGCTCAGGCGTTGGCTGGC 109783

Qy     944 GTGCCAATATTTACCGCCGACACCACCAAATCCATATTTGGATCGTTAATTAACGAGTGCGAGGC 1003
            |||                                                         |||
Db  109784 ACGGCAAATAATTACCGCCTAGGCCACAACATTGATTACCACAGCATCATCTGTCAGAAAACC 109843

Qy    1004 AATGTCGTAACAGTGCTTTTGGCATTTGAGGCTGTGATGGGGGATGGATTGGG 1053
            |||                                            |||
Db  109844 AGCGTGGTCACAGTGCGTCTTTTGCATTTGAGCCAGTAGTAAATCCGAACAATTGG 109893

Qy    2 AAAGCGTGACCACCATGGTTAACAAAACTTATCACCTCGCTCATTATATAATTTTAAATTGATCC 61
```

RESULT 8

LOCUS	AX932204 1314 bp DNA linear PAT 22-DEC-2003
DEFINITION	Sequence 157 from Patent WO03087353.
ACCESSION	AX932204
VERSION	AX932204.1 GI:40312625
KEYWORDS	.
SOURCE	Haemophilus influenzae
ORGANISM	Haemophilus influenzae

Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
Pasteurellaceae; Haemophilus.

REFERENCE   1

AUTHORS   Edwards,A., Dharamsi,A., Vedadi,M., Domagala,M., Houston,S.,
            Awrey,D., Beattie,B., Mansoury,K., Ouyang,H., Vallee,F.,
            Richards,D., Nethery,K., Virag,C., Buzadzija,K., Pinder,B.,
            Alam,M.Z., Tai,M., Canadien,V. and Kanagarajah,D.

JOURNAL   Patent: WO 03087353-A 157 23-OCT-2003;
            Affinium Pharmaceuticals, Inc. (CA)

FEATURES   Location/Qualifiers
     source   1..1314
                     /organism="Haemophilus influenzae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:727"

ORIGIN

Query Match   11.8%; Score 160.4; DB 2; Length 1314;
Best Local Similarity   49.6%; Pred. No. 1.2e-40;
Matches 683; Conservative   0; Mismatches 606; Indels   89; Gaps   7;

```
Qy      1 TAAACGGTGAACCATGTTAACAACTTATCACCTGGCTCATTATTAATTTTTAAATTGATC   60
Db   1308 TAAACATTGAGCTAAATGCGTAAATTCTTCGCGCCGCTTTTCAAAAGAAGCAAACTGATC 1249

Qy     61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA  120
Db   1248 GAGACTTGCACAAGCAGGCGACAATAATACCA---------------TATCTCCGCTTGCA 1201

Qy    121 TACTGCCCGATCAAGTGTACCAACTGATATATTTTTGATAGTCCAAATGTGCATCACCTTT  180
Db   1200 TGTTGGGCGGTAAAAAATTCTATCGCTTGTTCCATTGTATCGGAACAA-----ATAACTTTG 1147

Qy    181 GAGTAAATCTTGCTCAATTATTCGGGGCGGGCCATCGTAGCCTTTGACCGATTAAAATACACATCTCACCCA  240
Db   1146 CGATGAAAATTTGCAAGCAGCGGACACCATCTCGACCAAAACAATAAAATAATGTGTGG 1087

Qy    241 CTGCATAACATATGTGACAGCTCACTAAAATCTTGACACCTTTACCAAGCCCACCCAAAAT  300
Db   1086 TTGATTAATTAATTCAGCTAATTCTGAAAAAAATACAGCCCCTTTTTCCGTCCGTTAGCAA 1027

Qy    301 CAATGCCAGTGATGATAGCCATAAAACCGTCCCAAGCCCATCAATTGCTGAAATGGTTGA  360
Db   1026 CAAATGCAATTTACCCTCAATATA---------AAGCCCAGCCAATGCAGCGAACTGTACT  976

Qy    361 GCCGATGTTTGTACCTTTTGAAATCATTAAAAATATGCCTTGCTATCAATATCAATATA  420
Db    975 ... 420
```

```
RESULT 9
LOCUS       AX191762                11498 bp    DNA     linear   PAT 15-AUG-2001
DEFINITION  Sequence 44 from Patent WO0149775.
ACCESSION   AX191762
VERSION     AX191762.1  GI:15209931
KEYWORDS
SOURCE      Haemophilus influenzae
  ORGANISM  Haemophilus influenzae
            Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
            Pasteurellaceae; Haemophilus.
REFERENCE   1
  AUTHORS   Iversen,P.L.
  TITLE     Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 44 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..11498
                     /organism="Haemophilus influenzae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:727"
ORIGIN

Query Match      11.6%;  Score 160.4;  DB 2;  Length 11498;
Best Local Similarity   49.6%;  Pred. No. 1.9e-40;
Matches  683;  Conservative   0;  Mismatches  606;  Indels  89;  Gaps  7;
```

RESULT 10
BD426631_11/c
WPCOMMENT
Sequence split into 19 fragments   LOCUS BD426631 Accession BD426631

Fragment Name	Begin	End
BD426631_00	1	110000
BD426631_01	100001	210000
BD426631_02	200001	310000
BD426631_03	300001	410000
BD426631_04	400001	510000
BD426631_05	500001	610000
BD426631_06	600001	710000
BD426631_07	700001	810000
BD426631_08	800001	910000
BD426631_09	900001	1010000
BD426631_10	1000001	1110000
BD426631_11	1100001	1210000
BD426631_12	1200001	1310000
BD426631_13	1300001	1410000
BD426631_14	1400001	1510000
BD426631_15	1500001	1610000
BD426631_16	1600001	1710000
BD426631_17	1700001	1810000
BD426631_18	1800001	1830121

Continuation [12 of 19] of BD426631 from base 1100001 (BD426631 The Nucleotide Sequence

Query Match 11.6%;   Score 160.4;   DB 2;   Length 110000;
Best Local Similarity 49.6%;   Pred. No. 3e-40;
Matches 683;   Conservative 0;   Mismatches 606;   Indels 89;   Gaps 7;

```
Db  103428  ------------------GAATAATTCAAATATCGCCGATTACTTCCACTCCCGCTTTAA 103388
Qy  1140  TGGCGCGGGCACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGAGTGATGAGTCGTCGG 1199
Qy  103387  GTGCGGTTTGAATTTCTGGTGTTTTACCGCAAGCGCCTGGGCTAATACGAACATATATCGC 103328
Db  103387  GTGCGGTTTGAATTTCTGGTGTTTTACCGCAAGCGCCTGGGCTAATACGAACATATATCGC 103328
Qy  1200  CTGAGCTTAAGAGTTCACTGTCCAAAAATGACAAATGAACGCCTGTGGGCAGTT 1259
Db  103387  ...
```

RESULT 11
BD426631_12/c
WPCOMMENT
Sequence split into 19 fragments   LOCUS BD426631 Accession BD426631

Fragment Name   Begin   End
BD426631_00   1   110000
BD426631_01   100001   210000
BD426631_02   200001   310000
BD426631_03   300001   410000
BD426631_04   400001   510000
BD426631_05   500001   610000
BD426631_06   600001   710000
BD426631_07   700001   810000
BD426631_08   800001   910000
BD426631_09   900001   1010000
BD426631_10   1000001   1110000
BD426631_11   1100001   1210000
BD426631_12   1200001   1310000
BD426631_13   1300001   1410000
BD426631_14   1400001   1510000
BD426631_15   1500001   1610000
BD426631_16   1600001   1710000
BD426631_17   1700001   1810000
BD426631_18   1800001   1830121

Continuation [13 of 19] of BD426631 from base 1200001 (BD426631 The Nucleotide Sequence

Query Match   11.6%;   Score 160.4; DB 2; Length 110000;
Best Local Similarity 49.6%;   Pred. No. 3e-40;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

RESULT 12
AR274513_11/c
WPCOMMENT
Sequence split into 19 fragments   LOCUS AR274513 Accession AR274513

[Sequence alignment data continues - dense nucleotide alignment blocks]

Fragment Name	Begin	End
AR274513_00	1	110000
AR274513_01	100001	210000
AR274513_02	200001	310000
AR274513_03	300001	410000
AR274513_04	400001	510000
AR274513_05	500001	610000
AR274513_06	600001	710000
AR274513_07	700001	810000
AR274513_08	800001	910000
AR274513_09	900001	1010000
AR274513_10	1000001	1110000
AR274513_11	1100001	1210000
AR274513_12	1200001	1310000
AR274513_13	1300001	1410000
AR274513_14	1400001	1510000
AR274513_15	1500001	1610000
AR274513_16	1600001	1710000
AR274513_17	1700001	1810000
AR274513_18	1800001	1830121

Continuation (12 of 19) of AR274513 from base 1100001 (AR274513 Sequence 1 from patent U...

Query Match 11.6%;   Score 160.4;   DB 2;   Length 110000;
Best Local Similarity 49.6%;   Pred. No. 3e-40;   Mismatches 606;   Indels 89;   Gaps 7;
Matches 683;   Conservative 0;

```
Qy      1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTTCGCTCATTATATAATTTTTAAATTGATC 60
            |||||   |||  | ||||||||  ||   ||||||| |||||||| ||||||||||||||||
Db 104439 TAAAACATTGAGCTAAATGGGTAAAATTCTTCGCGGGCCTTTTTCAAAAGAAGCAAAACTGATC 104380

Qy     61 GAGACTTGCACAAGCAGGCGGACAATAATACCA----------TATCTCCGCTTTGCAA 120
            |||  ||||| || || ||| ||| ||||  ||          |||||||||||||||
Db 104379 CATGCTAGCACAAGCAGTGATCGACAAAAGACAGCAGCCCTTAGCGTGCTGTTGACTGGCTGCAA 104332

Qy    121 TACTGCCCGATCAAGTGTACCAACCTGATATATATTTTAGATGCATCACCTTT 180
            |||| |   ||||||  ||||  || || |  |||||||||||||||||||
Db 104331 TGTTGGGGGGTAAAAATTCTATCCGTTGTTCCATTGTATCGGAACAA------ATAACTTTG 104278

Qy    181 GAGTAAATCTTGCTCAATTATCGGGGCCATCTTGACCGATTAAAAATACACATATCCACCCA 240
            |||   ||||||||||| |  || ||||||| |||||  |  |||||||||||||||||||
Db 104277 CGATGAAAAATTTTTGCAAGGCAGGCGGCACCATCTCGACCAAAACAATAACATATAATGTGTGG 104218

Qy    241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCCACCCAAAAT 300
            |  |     ||||  |||  ||   | ||||||   ||||||||||||||  ||||||||||
Db 104217 TTGATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTCCGTCTCCGCCTAGCAA 104158

Qy    301 CAATGCCAGTGATGAATATGCCAAGCCCATCAAATTGCTGCAATGGTTGA 360
            |    |    || ||  ||  |  |||||| |   ||  |||||||||||
Db 104157 CAAATGCAAATTACCCTCAATATA------AAGCCCAGCCAATGCAGCAACTGTACT 104107

Qy    361 GCCGATGTTGTTGTACCTTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
            |||  | |||| ||||| ||||| |||| ||||||||||  ||||| |  |||||||
Db 104106 CCCCACATTTGTTGCTTTAGAGTCATTAATACGAATGCCATTAGCTTGATGCATCCATGGG 104047

Qy    421 TTCACACGGATGCGGGTAGCCCTTTGAAAGTCTTCAAGCATGGCATCCATGGG 480
            |||    || ||||||||  | ||||||||||  ||||||| ||||||||||
Db 104046 TTGAAACGATGACTAACCCTTTGAAAATGACGAAGTGCGGTACGAAGTAAATCTAAAATT 103987

Qy    481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
            |||||| || | | ||||||||| |||| ||||||||| |||||||||||||||||||||
Db 103986 AATACCTATAGCTTGTGTGCCAAATGCTGTTGCTGCCAAAATGTTCATATATAATAATTATGGCGACC 103927

Qy    541 TTTGCATTTTTAATTGATGGATTGGTAAAAG 600
            |||||| |||| ||| ||||| ||  |||||
Db 103926 AACCAAAGTAGCTTCTTCACAAAGTAAAATCACTTTTTTTACCATTAAAATATTGCTTT 103867

Qy    601 ATCACCATCATCAAATAAATAAAAAATCTGC-CGTTTGGTGGGCTAAGTAAGCTATTTGTTG 659
            ||   ||  ||||  |||||||||  |||||  ||||||||||| ||||| |||||||||
Db 103866 GCCATTTTCAGTTTTTAGCCAATAATCCGCACTATTTCCGCAAAAGAAAACGGTATGTTT 103807

Qy    660 TGATAAATGGCTGTGTGATTTGGGGCGCAACGGCTTGTCAGTGTGTGGTCAAGCCAATTGCCTTGGCTAAAT 719
            |||  ||||||| ||||||||||    |||||  |||||||||||| |||||||||
Db 103806 CGCCTTGATTTCGTTTTTCCCAAAAGTCAGCC------------------- 103775
```

RESULT 13
AR274513 12/c
WPCOMMENT
Sequence split into 19 fragments   LOCUS AR274513 Accession AR274513

Fragment Name	Begin	End
AR274513_00	1	110000
AR274513_01	100001	210000
AR274513_02	200001	310000
AR274513_03	300001	410000
AR274513_04	400001	510000
AR274513_05	500001	610000
AR274513_06	600001	710000
AR274513_07	700001	810000
AR274513_08	800001	910000
AR274513_09	900001	1010000
AR274513_10	1000001	1110000
AR274513_11	1100001	1210000
AR274513_12	1200001	1310000
AR274513_13	1300001	1410000
AR274513_14	1400001	1510000
AR274513_15	1500001	1610000
AR274513_16	1600001	1710000
AR274513_17	1700001	1810000
AR274513_18	1800001	1830121

Continuation (13 of 19) of AR274513 from base 1200001 (AR274513 Sequence 1 from patent U...

Query Match 11.6%;   Score 160.4;   DB 2;   Length 110000;

```
Qy   1020 TTTTGGCATTTGAGCCTGTTGTGGTGCGATGATGTTGGTTGTGTTTTGTTTTGAGTTAGATCTC 1079
            ||||      |||    |||| ||||| |||||| | ||  |||||||||||
Db 103480 TTTTTACCATTTGAACCTGTAATCCCCACAAATTGGCTTTGTCGGCTCGCGGGCA------- 103429

Qy   1080 GTGCTTTGAGTGTATCAATGAACAATTGAACATCCCTGATCACACGGAATGCCTTGAGCTT 1139
            ||||||||||| | || |||||||||||||| ||||| |  || ||||||||||||||
Db 103428 -------GAAATAATTCAATATATTCAAAATTGAACATCGCTTGGGGTTCCACTCCGTTTAA 103388

Qy   1140 TGGCGGCGACAATACTCGGGTTCTTGGGGTTTAATGCGGGGGTTGAGGAGATCTGATCGG 1199
            ||| ||  ||||||||||||| |||||| ||||||| ||||| |||| || || ||||
Db 103387 GTGCGGTTTGAATTTCTGGTGTGTTTTTTACCCCAAGCCCTGGGGTCTAATAACAATCATATCGC 103328

Qy   1200 CTGAGCTTAAGAGTTCACTGTCCCAAGCTGCCAAAATGACAAATGAAGCGCAGTT 1259
            ||||| |||||||  |||| |   ||||||  ||||||||||| |||||||||||
Db 103387 TTTCAAGTACCATTCCTGATTTAAACTACCAGTATGACGTAAATGTGAGGGAGGCAGTT 103268

Qy   1260 TATCTGCCAAGGTAGGATTGGGATTACCATCGATGAGATGCTAACTTTGTTGACCGGTGATTGA 1319
            || || || ||||| ||||| ||||||||||| | |||| |||||||||| |||||||
Db 103267 TATCAATACCAGTAGGATTTTTTTTCGAGTATCAATACACACAAAATCAACTGTTGGGGATA 103208

Qy   1320 CCAAAAAATTCACACAGCTGACAGTCCAGAGTCCCAAACGACGAGGGCATATTTTTT 1377
            ||       || ||| || ||||||||||| |||||  |||||||||||||||
Db 103207 AGAGATAATCCACACAGCAAGAGAAAGACCTGTTTTGCCAAGCCGATGATAGTAATATATTTT 103150
```

```
Best Local Similarity  49.6%;   Conservative   0;   Pred. No. 3e-40;   Mismatches 606;   Indels  89;   Gaps   7;
Matches 683;

Qy        1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTAAATTGATC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4439 TAAACATTGAGCTAAATGCCTAAATCTTCGCCGGCCTTTTCAAAGAGCAAAACTGATC 4380

Qy       61 CATGCTAGCACAACAGCAGGTGACAAAAGACACAGCCCTAGCGGGTGCTTTGACTGGTGCCAA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     4379 GAGACTTGCACAAGCAGGGGACAATAATAACCA------------TATCTCCGCTTTGCAA 4332

Qy      121 TACTGCCCGATCAAGTGTACCAACCTGATATATATTTTAGATGCCAAATGTCACACCTTT 180
            ||||||||||||||||||||||||||||||||||
Db     4331 TGTTGGGGCGTAAAAATTCTATCGCTTGCCATTGTATCGGAACAA------ATAACTTTG 4278

Qy      181 GAGTAAATCTTGCTCAATTATCGGGGCCATCTTGTTCCATTGTATCGGAACAA 240
            ||||||||||||||||||||||||||||||
Db     4277 CGATGAAAATTTTGCAAGCAGCGGCACCATTCGACGCAAAAATAACACTATCCACCCA 4218

Qy      241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTTACCCAAAAAT 300
            |||||||||||||||||||
Db     4217 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCCGTCTCCGCCTAGCAA 4158

Qy      301 CAATGCCAGTGATGAATATGCCATAAACCGTGCCAAGCCCATCAATTGCTCGCAATGGTTGA 360
            ||||||||||||||||
Db     4157 CAAATGCAATTACCCTCAATATA------------AAGCCCAGCCAAATGCAGCAACTGTACT 4107

Qy      361 GCCATGTTTTTGTTGTACCTTTTTTGAAATCATTGAAATATGCCTACTAATCATCAATAATA 420
            |||||||||||||||||
Db     4106 CCCCACATTTTGTTGTTTAGAGTGTAATTAATTCCAACAGAATCCCATTAGCTTGAATGCAGTA 4047

Qy      421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTTAAAGGGTTATCAAGCATGGCATCCATGGG 480
            ||||||||||||||||||
Db     4046 TTGAAAACGATGATCTAACCCTTTGAAAATGACGAAGTGCGGGTAGGAAGTTGAATCTAAATT 3987

Qy      481 AAGATTGGCAGCTTGTGCCTAAGGCAAGTGCGAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
            |||||||||||||||||||
Db     3986 AATACCTATAGCTTGTGTCGCAATGCTGTTGCTGCCAAAAGTTCATATATAATTCATCAACTGTTC 3927

Qy      541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTTTGATGATAAAATGCCATACC 600
            |||||||||||||||||
Db     3926 AACCAATGTAGCTTCTTCACACAGGTAAAATCACTTCATCTTTTACCATTAAAATATTTTCGG 3867

Qy      601 ATCACCATCATCAAATAATAAAAATCTGC--CGTTTGGGTCGGCTAAGTAAGCTATTTGTTG 659
            ||||||||||||||||||
Db     3866 GCCATTTTCAGTTTTTAGCCAATAAATCCCCACTATTTTTCCGCAAAAGAAAACGGTATGTTTT 3807

Qy      660 TGATAATGGTGTGATTTGGGGCAACGCTTGTCTCAGTGTGGTGCTCAAGCACATTGCTTGGCTAAAT 719
            |||||||||
Db     3806 CGCTTGATTTTCGTTTTCGTTTTCCCAAAAGTCAGCC-------------------- 3775

Qy      720 TGGCATCCGTCTTGGTGAGTGATGGCGGTTGGCAATTTGCCAATTTTTGTTG 779
            |||||||||||
Db     3774 ---TATCTTCATTGTTCAACACATGCCGGCACCTACTTAGCTTATGAAATGCGTAAATTGCTT 3718

Qy      780 CCAAATAATCCACCATGCCATCGTCGCCGAGCAGATGATCAGCGGAGCGGATTTTAAAATTG 839
            |||||||||
Db     3717 GGGCAATAATCTTCTAAAATCCATATAGCGGATCCATATGAATCTTCAGTTCAGTTCAAGACAG 3658

Qy      840 TTGCCCCTTGAGCTGGCCAAATTTGAGNATATGCTCAAGACTGACTGAAACTTGACGCAGCTCAAGCA 899
            ||||||||||||
Db     3657 TCGGTGCCGCGAGCTTTAAGCTAATAAGTGTCTCAAGCTGAAAAACTAGTA-----GGTA 3598

Qy      121 ...
```

Fragment Name	Begin	End
AR632719_00	1	110000
AR632719_01	100001	210000
AR632719_02	200001	310000
AR632719_03	300001	410000
AR632719_04	400001	510000
AR632719_05	500001	610000
AR632719_06	600001	710000
AR632719_07	700001	810000
AR632719_08	800001	910000
AR632719_09	900001	1010000
AR632719_10	1000001	1110000
AR632719_11	1100001	1210000
AR632719_12	1200001	1310000
AR632719_13	1300001	1410000
AR632719_14	1400001	1510000
AR632719_15	1500001	1610000
AR632719_16	1600001	1710000
AR632719_17	1700001	1810000
AR632719_18	1800001	1830121

Continuation (12 of 19) of AR632719 from base 1100001 (AR632719 Sequence 1 from patent US

LOCUS AR632719 Accession AR632719

Query Match   11.8%;   Score 160.4;   DB 2;   Length 110000;
Best Local Similarity 49.6%;   Conservative   0;   Pred. No. 3e-40;   Mismatches 606;   Indels 89;   Gaps   7;
Matches 683;

```
Qy        1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTAAATTGATC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   104439 TAAACATTGAGCTAAATGCCTAAATCTTCGCCGGCCTTTTCAAAGAGCAAAACTGATC 104380

Qy       61 CATGCTAGCACAACAGCAGGTGACAAAAGACACAGCCCTAGCGGGTGCTTTGACTGGTGCCAA 120
            ||||||||||||||||||||||||||||||||||||
Db   104379 GAGACTTGCACAAGCAGGGGACAATAATAACCA------------TATCTCCGCTTTGCAA 104332

Qy      121 TACTGCCCGATCAAGTGTACCAACCTGATATATATTTTAGATGCCAAATGTCACACCTTT 180
            ||||||||||||||||||||||
Db   104331 TGTTGGGGCGTAAAAATTCTATCGCTTGCCATTGTATCGGAACAA------ATAACTTTG 104278

Qy      181 GAGTAAATCTTGCTCAATTATCGGGGCCATCTTGTTCCATTGTATCGGAACAA 240
            ||||||||||||||||||
Db   104277 CGATGAAAATTTTGCAAGCAGCGGCACCATTCGACGCAAAAATAACACTATCCACCCA 104218

Qy      241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTTACCCAAAAAT 300
            ||||||||||||
Db   104217 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCCGTCTCCGCCTAGCAA 104158
```

```
RESULT 15
AR632719
WPCOMMENT
AR632719 12/c
Sequence split into 19 fragments    LOCUS AR632719 Accession AR632719
Fragment Name       Begin       End
AR632719_00         1           110000
AR632719_01         100001      210000
AR632719_02         200001      310000
AR632719_03         300001      410000
AR632719_04         400001      510000
AR632719_05         500001      610000
AR632719_06         600001      710000
AR632719_07         700001      810000
AR632719_08         800001      910000
AR632719_09         900001      1010000
AR632719_10         1000001     1110000
AR632719_11         1100001     1210000
AR632719_12         1200001     1310000
AR632719_13         1300001     1410000
AR632719_14         1400001     1510000
AR632719_15         1500001     1610000
AR632719_16         1600001     1710000
AR632719_17         1700001     1810000
AR632719_18         1800001     1830121

Continuation (13 of 19) of AR632719 from base 1200001 (AR632719 Sequence 1 from patent US

Query Match      11.6%;    Score 160.4;   DB 2;    Length 110000;
Best Local Similarity   49.6%;    Pred. No. 3e-40;
Matches  683;    Conservative  0;   Mismatches 606;   Indels  89;   Gaps  7;
```

Left alignment column (Qy / Db):

Qy start	Qy end	Db start	Db end
301	360	104157	104107
361	420	104106	104047
421	480	104046	103987
481	540	103986	103927
541	600	103926	103867
601	659	103866	103807
660	719	103806	103775
720	779	103774	103718
780	839	103717	103658
840	899	103657	103598
900	959	103597	103541
960	1019	103540	103481
1020	1079	103480	103429
1080	1139	103428	103388
1140	1199	103387	103328
1200	1259	103327	103268
1260	1319	103267	103208
1320	1377	103207	103150

Right alignment column (Qy / Db):

Qy start	Qy end	Db start	Db end
1	60	4439	4380
61	120	4379	4332
121	180	4331	4278
181	240	4277	4218
241	300	4217	4158
301	360	4157	4107
361	420	4106	4047
421	480	4046	3987
481	540	3986	3927
541	600	3926	3867
601	659	3866	3807

```
Qy     660  TGATAATGGGTGTGATTTGGGGCAACGGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAAT  719
Db    3806  CGCTTGATTTTCGTTTTCCCAAAAGTCAGCC-------------------------------  3775

Qy     720  TGGCATCGTCTTGGTAGATGATGGGCGGTTTGGCAATTTCAAAAATACGCAAATTTTTGTG  779
Db    3774  ---TATCTTCATTGTTCAACACACTTACTTTAGCATTATGGATTAAATGCGTAATTTTGCTT  3718

Qy     780  CCAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGGAGAGATTTAAAATTG  839
Db    3717  GGGCGGATAAATCTTCTAAAATCCATATAGCGATCCATATGATCTTCAGTCACGTTCAAGACAG  3658

Qy     840  TTGCCCCTTGAGCTGCCAAAATTTGAGATAATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA  899
Db    3657  TCGCTGCCGCGAGCTTTTTAAGCTTATAAGTTGTCTCAAGCTGAAAAACTCTAGTA  3598

Qy     900  CCACCAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGGTGCCAATATTACCGC  959
Db    3597  CATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAATATTTCCGC  3541

Qy     960  CGACACCAACAAATCATGCCTGCATGTTTTTGCCATCTCGCCTACTAAATGTCGTAACAGTGC  1019
Db    3540  CCATACCAACTTTCACACCAGCACGCTTTCGCCCATTTCATAAACTAAAGTAGTTACCGGTAC  3481

Qy    1020  TTTTTGGCATTTGAGCCTGTGATGGGGATGATTGGTGTGTTGTTTGCTTGGAGTTAGATCTC  1079
Db    3480  TTTTTTACCATTTGAACCTGTAATCCCCACAATTGGCTTTGTCGCTGCGGGCA---------  3429

Qy    1080  GTGCTTTGAGTGTATCAAATGAAACAAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTT  1139
Db    3428  --------------GAATAATTCAATATCGCCGATTACTTCCACTCCCGCTTTAA  3388

Qy    1140  TGGCGGCGACAAATACTTCGGCGGTTCTTGGGGTTTAATGCCGGGGGGCTGATGAATCTGATCGG  1199
Db    3387  GTGCGGGTTTGAATTTCTGGTGTTTTTTTACCGGCAAGCCCTAATAACAATCATATCGC  3328

Qy    1200  CTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTT  1259
Db    3327  TTTCAAGTAACCATTCCCTGATTTAAACTACCAGTATGAAGAGGGATATTTTTGAGGAAGTT  3268

Qy    1260  TATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGTCTAACTTTGTGACCGTGATTGA  1319
Db    3267  TATCAATACCAGTAGGATTTTTTCGAGTATCAAATCACACGAATTATTAGCCTGTTGGGATA  3208

Qy    1320  CCAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAAACCGACGCGGCATATTTTTT  1377
Db    3207  AGAGATAATCCACACAAGAAAGACCTGTTTTGCCAAGCCCGATGATTAGTAATATATTTTTT  3150
```

QM nucleic - nucleic search, using sw model

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: May 12, 2006, 07:12:03 ; Search time 905 Seconds
                 (without alignments)
                 10162.736 Million cell updates/sec

Title: US-10-672-787-35_COPY_11357_12736
Perfect score: 1380
Sequence: 1 taaagcgtgaaccatcgttaa..........acgacggcatattctttga 1380

Scoring table: IDENTITY_NUC
                 Gapop 10.0, Gapext 1.0
Searched: 4996997 seqs, 333234630B residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   N_Geneseq_21:*
    1: geneseqn1980s:*
    2: geneseqn1990s:*
    3: geneseqn2000s:*
    4: geneseqn2001as:*
    5: geneseqn2001bs:*
    6: geneseqn2002as:*
    7: geneseqn2002bs:*
    8: geneseqn2003as:*
    9: geneseqn2003bs:*
   10: geneseqn2003cs:*
   11: geneseqn2003ds:*
   12: geneseqn2004as:*
   13: geneseqn2004bs:*
   14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	1380	100.0	1398	8	ACA39325	Aca39325 Prokaryot
c 2	1380	100.0	96109	4	AAF28548	Aaf28548 Genomic f
c 3	1370.4	99.3	1410	12	ADL02987	Adl02987 DNA encod
c 4	231	16.7	1344	8	ACA21070	Aca21070 Prokaryot
c 5	226.2	16.4	1416	9	ADA29995	Ada29995 DNA encod
c 6	160.4	11.6	1314	4	AAS53453	Aas53453 Haemophil
c 7	160.4	11.6	1314	8	ACA34366	Aca34366 Prokaryot
c 8	160.4	11.6	1314	11	ADL46439	Adl46439 aspartate
c 9	160.4	11.6	1314	14	AEC10868	Aec10868 Haemophil
c 10	160.4	11.6	110000	2	AAT42063_11	Continuation (12 o
c 11	160.4	11.6	110000	14	AEB39175_30	Continuation (13 o
c 12	157.6	11.4	110000	2	AAT42063_12	Continuation (31 o
c 13	157.6	11.4	14324	13	ADT05533	Adt05533 Haemophil
c 14	154.4	11.2	349980	13	ADT05649	Adt05649 Haemophil
c 15	154.4	11.2	14324	13	ADT05649	Adt05649 Haemophil
c 16	152.8	11.1	1281	8	ACA37295	Aca37295 Prokaryot
c 17	152.8	11.1	57367	14	AEB39164	Aeb39164 L. pneumo
c 18	152.8	11.1	110000	14	AEB42737_15	Continuation (16 o
c 19	152.4	11.0	1314	11	ADL46441	Adl46441 aspartate
c 20	152.4	11.0	1314	14	AEC10870	Aec10870 Haemophil
c 21	145.2	10.5	229301	14	AEB35719	Aeb35719 L. pneumo
c 22	142.2	10.3	1305	8	ACA42843	Aca42843 Prokaryot
c 23	139.6	10.1	1317	5	AAS52295	Aas52295 E. coli D
c 24	139.6	10.1	1317	4	AAH81460	Aah81460 Escherich
c 25	139.6	10.1	1317	8	ACA32359	Aca32359 Prokaryot
c 26	125.2	9.1	1323	10	ADF02545	Adf02545 Bacterial
c 27	123.6	9.0	1308	8	ACA44759	Aca44759 Prokaryot
c 28	117.8	8.5	1182	8	ACA45658	Aca45658 Prokaryot
c 29	117	8.5	1347	4	AAS54275	Aas54275 Pseudomon
c 30	117	8.5	1347	8	ACA42598	Aca42598 Prokaryot
c 31	117	8.5	1371	11	ABD09098	Abd09098 Pseudomon
c 32	117	8.5	1401	10	ABD09257	Abd09257 Pseudomon
c 33	113	8.2	1450	8	AAZ47131	Aaz47131 Pseudomon
c 34	113	8.2	1404	8	ACA53404	Aca53404 Prokaryot
c 35	111	8.0	1317	8	ACA51133	Aca51133 Prokaryot
c 36	110.8	8.0	4942	5	AAS89816	Aas89816 DNA encod
c 37	110.8	8.0	4944	5	AAS89273	Aas89273 DNA encod
c 38	110.8	8.0	4944	5	AAS90181	Aas90181 DNA encod
c 39	110	8.0	1350	8	ACA43846	Aca43846 Prokaryot
c 40	109.6	7.9	1311	10	ACF71682	Acf71682 Photorhab
c 41	109.6	7.9	110000	10	ACF67367_48	Continuation (49 o
c 42	109.6	7.9	110000	10	ACF65387_0	Acf65387 Photorhab
c 43	109.4	7.9	1317	4	AAS56341	Aas56341 Salmonell
c 44	109.4	7.9	1317	8	ACA51965	Aca51965 Prokaryot
c 45	103.8	7.5	1314	8	ACA35530	Aca35530 Prokaryot

ALIGNMENTS

RESULT 1
ACA39325
ACA39325/c
ACA39325 standard; DNA; 1398 BP.
XX
ID   ACA39325 standard; DNA; 1398 BP.
XX
AC   ACA39325;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Prokaryotic essential gene #20982.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Moraxella catarrhalis.
XX
XX
PN   WO200227183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT) ELITRA PHARM INC.
XX
PI   Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI   Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU35455.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 27195; 1766pp; English.
CC
CC   The invention relates to an isolated nucleic acid comprising any one of

```
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1398 BP; 408 A; 306 C; 295 G; 389 T; 0 U; 0 Other;

Query Match             100.0%;  Score 1380;  DB 8;  Length 1398;
Best Local Similarity   100.0%;
Matches 1380;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1383 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 1324

Qy     61 CATGCTAGCACACAAGCAGGTGACAAAAAGCACAGCCCTAGCCGTGCTTGACTGGCTGCCAA 120
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1323 CATGCTAGCACACAAGCAGGTGACAAAAAGCACAGCCCTAGCCGTGCTTGACTGGCTGCCAA 1264

Qy    121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAATGTGCATCACCTTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1263 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAATGTGCATCACCTTT 1204

Qy    181 GAGTAAATCTTGCTCAATTATTCGGGTGACGATTAAAATACACACTATCCACCCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1203 GAGTAAATCTTGCTCAATTATTCGGGTGACGATTAAAATACACACTATCCACCCA 1144

Qy    241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1143 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 1084

Qy    301 CAATGCCAGTGATGATATGCCATAAACGTGCCAAGCCCATCAATTGCCAATGGTTGA 360
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1083 CAATGCCAGTGATGATATGCCATAAACGTGCCAAGCCCATCAATTGCCAATGGTTGA 1024

Qy    361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCGATCATCAATAATATA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1023 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCGATCATCAATAATATA 964

Qy    421 TTCACAGCGGATGGGGTAGCCCTTGAAGAAGTCTTAAGGGTATCAAGCATGGCATCCATGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    963 TTCACAGCGGATGGGGTAGCCCTTGAAGAAGTCTTAAGGGTATCAAGCATGGCATCCATGG 904

Qy    481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    903 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 844

Qy    541 TTTGATTTTAAATTGATGAGGATTGGTAAAAGTTTTGTTTTTTGATGATAAAATGCATACC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    843 TTTGATTTTAAATTGATGAGGATTGGTAAAAGTTTTGTTTTTTGATGATAAAATGCATACC 784

Qy    601 ATCACCATCATCAAAATAAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGTCTATTTGTGT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    783 ATCACCATCATCAAAATAAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGTCTATTTGTGT 724

Qy    661 GATAATGGTGTGCATTTGGGGCGCAAGCGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 720
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    723 GATAATGGTGTGCATTTGGGGCGCAAGCGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 664

Qy    721 GGCATCGTTCTTGGTAGATGATGGGGGCGGTTTGGCGAATTTTCAAAAAATACGCAATTTTTGTGC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    663 GGCATCGTTCTTGGTAGATGATGGGGGCGGTTTGGCGAATTTTCAAAAAATACGCAATTTTTGTGC 604

Qy    781 CAAATAATTCACCATGCCATCGTGCCGATCAAGATCAGGCGGAGGAGATTTAAAATTGT 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 CAAATAATTCACCATGCCATCGTGCCGATCAAGATCAGGCGGAGGAGATTTAAAATTGT 544

Qy    841 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAACTTGACAGCTCAAGCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAACTTGACAGCTCAAGCAC 484

Qy    901 CACCAAAATCCATATTTTGGATCGTTAATATTCAAGTGCGCAGGCGGTGCCAATATTACCGCC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 CACCAAAATCCATATTTTGGATCGTTAATATTCAAGTGCGCAGGCGGTGCCAATATTACCGCC 424

Qy    961 GACACCCAACAATCATGCCCTGCCATCCTCGCCATCCTGCCCTCTAAATGTCGTGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 GACACCCAACAATCATGCCCTGCCATCCTCGCCATCCTGCCCTCTAAATGTCGTGCT 364

Qy   1021 TTTGGCATTTGAGCCTGTGATGGCCGATGAGTGGTGTTTTGCTTGGTTGGTTAGATCTCG 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 TTTGGCATTTGAGCCTGTGATGGCCGATGAGTGGTGTTTTGCTTGGTTGGTTAGATCTCG 304

Qy   1081 TGCTTTGAGTGTATCAAATGAACAAATTGAACATCGCTGATCACGAAATGCCTTGAGCTTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 TGCTTTGAGTGTATCAAATGAACAAATTGAACATCGCTGATCACGAAATGCCTTGAGCTTT 244

Qy   1141 GGCGCGCGACAATACTCGGGCGTTCTTCTTGGGTTAAATGCCGGGCGTGATGATCTCGAGC 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 GGCGCGCGACAATACTCGGGCGTTCTTCTTGGGTTAAATGCCGGGCGTGATGATCTCGAGC 184

Qy   1201 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAACGGCCCTGTGGGCAGTTT 1260
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAACGGCCCTGTGGGCAGTTT 124

Qy   1261 ATCTCCCAAGGTAGAGAGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGAGAC 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 ATCTCCCAAGGTAGAGAGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGAGAC 64

Qy   1321 CAAAAAAATTCACAGCTGACAGTCAGAGCTGCCCAAACCGACGACGGCCACATTTTTTTTGA 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 CAAAAAAATTCACAGCTGACAGTCAGAGCTGCCCAAACCGACGACGGCCACATTTTTTTTGA 4
```

RESULT 2
AAF28548

```
ID   AAF28548 standard; DNA; 96109 BP.
XX
AC   AAF28548;
XX
DT   04-APR-2001 (first entry)
XX
DE   Genomic fragment #35.
XX
KW   Genomic library; bacteria; human upper airway; otitis media; sinusitis;
KW   bronchopulmonary; endocarditis; meningitis; ss.
XX
OS   Moraxella catarrhalis.
XX
PN   WO20007968-A2.
```

XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US016649.
XX
PR   18-JUN-1999; 99US-0140121P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lagace RE, Patterson C, Berg KL;
XX
DR   WPI; 2001-041427/05.
XX
PT   Genomic library for identifying diagnostic and therapeutic compositions,
PT   and for identifying virulence factors, regulatory elements and drug
PT   targets, comprises Moraxella catarrhalis nucleic acids.
XX
PS   Claim 1; Page 345-368; 545pp; English.
XX
CC   The present invention relates to a Moraxella catarrhalis genomic library
CC   comprising of a combination of 41 nucleic acid molecules (see AAP28514-
CC   AAP28554). The library has a number of uses described in the
CC   specification e.g. is useful for aidentifying diagnostic and therapeutic
CC   compositions. M. catarrhalis (Brahhamella catarrhalis) is a large
CC   aerobic, gram-negative diplococcus, normally found among the bacterial
CC   flora of human upper airways. M. catarrhalis is known to cause acute,
CC   localised infections such as otitis media, sinusitis and bronchopulmonary
CC   infection and life-threatening, systemic diseases including endocarditis
CC   and meningitis
XX
SQ   Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 U; 0 Other;

Query Match   100.0%; Score 1380; DB 4; Length 96109;
Best Local Similarity   100.0%;
Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 TAAAGCGTGAACCAATGTTAACAAAACTTATCACCTCGGCTCATTATAATTTTAAATTGATC 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11357 TAAAGCGTGAACCAATGTTAACAAAACTTATCACCTCGGCTCATTATAATTTTAAATTGATC 11416

Qy   61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCCGGTGCTTGACTGGCTGCCAA 120
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11417 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCCGGTGCTTGACTGGCTGCCAA 11476

Qy   121 TACTGCCCGATCAAGTGTACCAACCTGATATATATTTAGATGCCAAAATGTGCATCACCTTT 180
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11477 TACTGCCCGATCAAGTGTACCAACCTGATATATATTTAGATGCCAAAATGTGCATCACCTTT 11536

Qy   181 GAGTAAAATCTTGCTCAAATTATCGGGGGCATCTTGACGGATTAAAATACACTATCATCCCA 240
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11537 GAGTAAAATCTTGCTCAAATTATCGGGGGCATCTTGACGGATTAAAATACACTATCATCCCA 11596

Qy   241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11597 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 11656

Qy   301 CAATGCCAGTGATGATGATGCCATAAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11657 CAATGCCAGTGATGATGATGCCATAAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 11716

Qy   361 GCCGATGTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCTATCAATATCAATAATA 420
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11717 GCCGATGTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCTATCAATATCAATAATA 11776

Qy   421 TTCACAGCGGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 480
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11777 TTCACAGCGGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 11836

Qy   481 AAGATTGGCAGCTGTGCCTAAGGCAGCATTGAGTAGGTTATGTCTTGCC 540
             |||||||||||||||||||||||||||||||||||||||||||||||||
Db   11837 AAGATTGGCAGCTGTGCCTAAGGCAGCATTGAGTAGGTTATGTCTTGCC 11896

Qy   541 TTTGATTTTTAATTGATTGGATTGGTAAAAGTTTGTTTTTTTGATGATATAAATGCCATACC 600
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11897 TTTGATTTTTAATTGATTGGATTGGTAAAAGTTTGTTTTTTTGATGATATAAATGCCATACC 11956

Qy   601 ATCACCATCATCAAAATAAAATAAAAATCTGCCGTTTGCCCGTTTGGCTAAGTAAGCTATTTGTGT 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11957 ATCACCATCATCAAAATAAAATAAAAATCTGCCGTTTGCCCGTTTGGCTAAGTAAGCTATTTGTGT 12016

Qy   661 GATAATGGTGTGATTTGGGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 720
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12017 GATAATGGTGTGATTTGGGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 12076

Qy   721 GGCATCCGTCTTCGGTAGATGATGGCGGTTTGGCAATTTCAAAAAATACGCCAATTTTTGTGC 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12077 GGCATCCGTCTTCGGTAGATGATGGCGGTTTGGCAATTTCAAAAAATACGCCAATTTTTGTGC 12136

Qy   781 CAAATAAATCCACCATAGCCATCCTGCCCATCAAGAATGAGCAGGTTTAAAAATTGT 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12137 CAAATAAATCCACCATAGCCATCCTGCCCATCAAGAATGAGCAGGTTTAAAAATTGT 12196

Qy   841 TGCCCCTTGAGCGTGCCAAATTTGAGATATGCTTCAAGCTGAAAAACTTGACAGCTCAAGCAC 900
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12197 TGCCCCTTGAGCGTGCCAAATTTGAGATATGCTTCAAGCTGAAAAACTTGACAGCTCAAGCAC 12256

Qy   901 CACCAAAATCCATATTTTTGGATCGTTAATAATTCAAGTGCAGCAGCGGTGCCAATATTACCGCC 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12257 CACCAAAATCCATATTTTTGGATCGTTAATAATTCAAGTGCAGCAGCGGTGCCAATATTACCGCC 12316

Qy   961 GACACCAACAATCATGCCTCGCATGTTTTTGCCATCTCGCCTACTAAATCGTTAACACGTGCT 1020
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12317 GACACCAACAATCATGCCTCGCATGTTTTTGCCATCTCGCCTACTAAATCGTTAACACAGTGCT 12376

Qy   1021 TTTGGCATTTGAGCCTGTGAATGGCGATGATGGTGTGTTTGTTGCTTGAGTTAGATCTCG 1080
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12377 TTTGGCATTTGAGCCTGTGATGGCGATGATGGTGTGTTTGTTGCTTGAGTTAGATCTCG 12436

Qy   1081 TGCTTTGAGTGTATCAATGAACAATTCGACACATCGCTCATCACAGGAATGCCTTGAGCTTT 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12437 TGCTTTGAGTGTATCAATGAACAATTCGACACATCGCTCATCACAGGAATGCCTTGAGCTTT 12496

Qy   1141 GGGCGGGGCACAATACTCGGCGGTTCTTCGGGTTAAATGCCGGGGCGTGATCGATCTGAGCGGC 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12497 GGGCGGGCACAATACTCGGCGGTTCTTCGGGTTAAATGCCGGGGCGTGATGATCTGATCGGC 12556

Qy   1201 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAAATGACAAATGAAACGCCTGTGGGCAGTTT 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12557 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAAATGACAAATGAAACGCCTGTGGGCAGTTT 12616

Qy   1261 ATCTGCCAAGGTAGGATGGATTGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGAC 1320
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12617 ATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGAC 12676

Qy   1321 CAAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCACATATTTTTTGA 1380
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12677 CAAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCATATTTTTTGA 12736

RESULT 3
ADL02987/c
ID   ADL02987 standard; DNA; 1410 BP.
XX
AC   ADL02987;
XX
DT   06-MAY-2004 (first entry)
XX
DE   DNA encoding a M. catarrhalis protein #673.
XX
KW   ds; gene; Moraxella catarrhalis; infection.
XX
OS   Moraxella catarrhalis.
XX
PN   US6673910-B1.
XX
PD   06-JAN-2004.

```
XX
PF   04-APR-2000; 2000US-00540236.
XX
PR   08-APR-1999; 99US-0128416P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Breton GL;
XX
DR   WPI; 2004-178127/17.
DR   P-PSDB; ADL04907.
XX
PT   New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT   preparing a composition for diagnosing, preventing or treating infection
PT   caused by Moraxella catarrhalis.
XX
PS   Disclosure; SEQ ID NO 673; 499pp; English.
XX
CC   The invention relates to an isolated nucleic acid encoding an Moraxella
CC   catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC   composition for diagnosing, preventing or treating infection caused by
CC   Moraxella catarrhalis. The present sequence represents DNA encoding a M.
CC   catarrhalis protein.
XX
SQ   Sequence 1410 BP; 409 A; 307 C; 302 G; 392 T; 0 U; 0 Other;

Query Match     99.3%;  Score 1370.4;  DB 12;  Length 1410;
Best Local Similarity   99.6%;
Matches 1374;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;  Pred. No. 0;

Qy          1 TAAAGGCTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATATAATTTTTAAAATTGATC 60
              ||||||| |||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db       1392 TAAAGGGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATATAATTTTTAAAATTGATC 1333

Qy         61 CATGCTAGACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTGACTGGCTGCCAA 120
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1332 CATGCTAGACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTGACTGGCTGCCAA 1273

Qy        121 TACTGCCCGATCAAGTGTACCAACCTGATATTTAGATGCCAAAATGTGCATCACCTTT 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1272 TACTGCCCGATCAAGTGTACCAACCTGATATTTAGATGCCAAAATGTGCATCACCTTT 1213

Qy        181 GAGTAAATCTTGCTCAATTATCGGGGCCATCTGATAATAAATACACTATCCACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1212 GAGTAAATCTTGCTCAATTATCGGGGCCATCTGATAATAAATACACTATCCACCCA 1153

Qy        241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCAAAAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1152 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCAAAAAT 1093

Qy        301 CAATGCCAGTGATGATATGCCATAAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1092 CAATGCCAGTGATGATATGCCATAAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 1033

Qy        361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCCATCATCAATTCATCAATATA 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1032 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCCATCATCAATTCATCAATATA 973

Qy        421 TTCACAGCCGATGGGGTAGCCCCTTTGAAAGTTCTTAAGGGGTATCAAGCATGGCATCCATGGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        972 TTCACAGCCGATGGGGTAGCCCCTTTGAAAGTTCTTAAGGGGTATCAAGCATGGCATCCATGGG 913

Qy        481 AAGATTGGCAGCTGTGTGCCTAAGCAGAGTAAAAGGCATTGAGTAGGTTATGCCATACC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 AAGATTGGCAGCTGTGCCTAAGCAGAGTAAAAGGCATTGAGTAGGTTATGCCATACC 853

Qy        541 TTTGATTTTTTAATTGATGGATTGGTTGTTTTTTTTGATGATAAATGCCATACC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        852 TTTGATTTTTTAATTGATGGATTGGTTGTTTTTTTTGATGATAAATGCCATACC 793

Qy        601 ATCACCATCATCAAATAAATAACTCTCCGTTGGCTAAGTAAGCTATTTGTTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 ATCACCATCATCAAATAAATAAAATCTGCCGTTGGCTAAGTAAGCTATTTGTTGT 733

Qy        661 GATAATGGTCGTTGATTTGGGGGCAACGCGTTGTCAGTGCGTGGGTCAAGCAGCATTGCTTGGCTAAATT 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        732 GATAATGGTCGTTGATTTGGGGGCAACGCGTTGTCAGTGCGTGGGTCAAGCAGCATTGCTTGGCTAAATT 673

Qy        721 GGCATCGTCTTGGCTAGATGATGGGGCGGTTTGGCAATTTCAAAATACGCAAATTTTTGTGC 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        672 GGCATCGTCTTGGCAGATGATGGGGCGGTTTGGCAATTTCAAAATACGCAAATTTTTGTGC 613

Qy        781 CAAATAAATCCACCATGCCATCGTGCCGATCAAGAATGATCAGCGGAGAGATTAAAATTGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        612 CAAATAAATCCACCATGCCATCGTGCCGATCAAGAATGATCAGCGGAGAGATTAAAATTGT 553

Qy        841 TGCCCCTTGAGCTGCCAAAATTTGACATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCAC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        552 TGCCCCTTGAGCTGCCAAAATTTGACATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCAC 493

Qy        901 CACCCAAAATCCATATTTTTTGGATCGTTAAATAATTCAAGCTCAAGCTGAAAAACTTGACAGCTCAAGCAC 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        492 CACCCAAAATCCATATTTTTTGGATCGTTAAATAATTCAAGCTGCCAATAATTATTACCGCC 433

Qy        961 GACACCCAACAATCAATGCCTCCATGTTTTGCCATCCTGCGCCCTACTAATGTCGTAACAGTGCT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        432 GACACCCAACAATCAATGCCCTGCAGTTTTTGCCATCCTGCGCCCTACTAATGTCGTAACAGTGCT 373

Qy       1021 TTTGGCATTTGAGCCTGTGATGGGGCGATTGGGTCGTGTTTTGCTTGAGTTGAGTTAGAATCTCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        372 TTTGGCATTTGAGCCTGTGATGGGGCGATTGGGTCGTGTTTTGCTTGAGTTGAGTTAGAATCTCG 313

Qy       1081 TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGGAATGCCTTGAGCCTTT 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        312 TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGGATGCCTTGAGCCTTT 253

Qy       1141 GGCGGCGACAATACTCGGGCGTTCTTGGGGTTAAGTCGTGGGGGCGCCGCTGATCGATCGGC 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        252 GGCGGCGACAATACTCGGGCGTTCTTGGGGTTAAGTCGTGGGGGCGCCGCTGATCGATCGGC 193

Qy       1201 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAAATGCATCGATGATGACCAGTCAGTTT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAAATGCATCGATGATGACCAGTCAGTTT 133

Qy       1261 ATCTGCCAAGGTAGGAATTGGGGATTACCATGATGATGATGCTAACTTTGTGACCGTGATTGAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 ATCTGCCAAGGTAGGAATTGGGGATTACCATGATGATGATGCTAACTTTGTGACCGTGATTGAC 73

Qy       1321 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAAACCGACGAGGGCATTATTTTTTTGA 1380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAAACCGACGAGGGCATTATTTTTTTGA 13
```

```
RESULT 4
ACA21070/c
ID   ACA21070 standard; DNA; 1344 BP.
XX
AC   ACA21070;
XX
DT   19-JUN-2003  (first entry)
XX
DE   Prokaryotic essential gene #2727.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Acinetobacter baumannii.
XX
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.
PR   06-SEP-2001; 2001US-00948993.
```

```
PR   25-OCT-2001; 2001US-0342923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-0362699P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI   Wall D, Trawick JD, Carr GJ, Yanamoto R, Forsyth RA, Xu HH;
XX
DR   WPI; 2003-029926/02.
DR   P-PSDB; ABU17200.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 8940; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1344 BP; 385 A; 254 C; 318 G; 387 T; 0 U; 0 Other;
```

```
Query Match            1344 BP;
Best Local Similarity   16.7%;    Score 231;   DB 8;   Length 1344;
Matches   721;  Conservative  0;  Mismatches 590;  Indels 66;  Gaps 7;
          52.4%;            Pred. No. 9.6e-56;
```

```
Qy      2 AAAGGGTGAACCANCATTGTTAACAAAACTTATCACCTCGCTCATTATAAATTAANATTTTAAATTGNATCC 61
Db   1340 AACGAATTGACGCACGCAACAACTCGTCACCACGGTCATTATATAACTTTTAAACATATTATCA 1281

Qy     62 ATGTTAGCACAAGCAGGTGACAAAAGCACACCCTAGCGCGTGTCATTATTAACTTTTAAACATATTATCA 121
Db   1280 AAACTTGCACATGCTGGTGAT------AGCAATACCACATCTTCAGCTTGTGTTTCAC 1229

Qy    122 ACTGCCGGATCAAGTGTACCAACCTGATATATTTAGATGCCAAARGTGCATCACCTTTG 181
Db   1228 GTTGACACAGCTCGACGCTTCTTTAAGCGTTGCTGCCATGTAAAATTTTACTTGGGGCT 1170

Qy    182 AGTAAATCTTGCTCAATTATCGGGGGCCATCTTGACCGATTAAATAACACACTATCCACCAC 241
Db   1169 TGAATGGGCTTGCTCGATTGACCGGCGCCATCTTCACCAATCAATTACCAACTTTGGCATAT 1110

Qy    242 TGCATAACACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAATC 301
Db   1109 TTTTCGGATAGAAGCAGTAAAGGAGGAGAAAAATCTTGCCCTTGCCCTGTCCACTTAAAATA 1050

Qy    302 AATGCCA------GTGATGAATATGCCATAAAACCGTGCCAAGAGCCCATCAAATTGCTGCAAGT 355
Db   1049 AGTGCAAACCTTACCTTTTTTCACTTCAARGGACGAGCAGCCACTTAAGCCCATCAAATTGCTGCAAGT 990

Qy    356 GTTGAGCCGGATGGTTGTACCTTTTTGAATCATTAAAATATGCCTTGCGTATCAAATATCATCA 415
Db    989 ATACCTTGAATATATAAAAATCCGAGCTTTTAAATTAACCCTGTAAAACCACGACGAAGCCAA 930

Qy    416 ATATATTCACAGCCATGCGGTACGCGTAGCCCCTTTGAAAGTCTTAAGGGTATCAAGGCATGGCATCC 475
Db    929 ACGTACTCACAGGGATGCTCTAATCCTTTAAATCGTTTTAAAGTTTCAAGCATTGACTCC 870

Qy    476 ATGGGAAGCATTGGCGAGCTGTGCCTAAGCGCAAGTGCAGATAAGGCCATTCAGTAGTTATGC 535
Db    869 GTAGCACCAAACATTGGGTACCTTTAGAGTGTCATTATAATAAACGGACATCACGCACAGTTTTA 930

Qy    536 TTGCCTTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTGATGATAATAAAATGCC 595
Db    809 ATACCTTGAAATATATAAAAATCCGAGCTTTTAAATTAACCCTGTAAAACCACGACGAAGCCAA 750

Qy    596 ATACCATCACCATCATCAAATAAATAAAATCTGCCGGTTTGGTGGCTAAGTAAGCTATTT 655
Db    749 ACGGTACGGTCGGCATCTCTTAAAACGCCCATATTGATT 698

Qy    656 GTTTGTGATAATGGTGTGATTTGGGGCAACGCTTGTCAGCTGGTGGTCAAGCATTGCTTGGCT 715
Db    697 TTAAACCAAAGCTTTGCATTGGTGTCGTTATCGGAACAAAGTGGCGGCTTA------ 647

Qy    716 AAAATTGGCATCGTCGTTCGGTTAGATGATGGCGGTTGGCAATTTCAAAAAATACGGCAAATTTT 775
Db    646 ----AACCATCGTCTCGGTTAAATACAACTTTTTAGCGCCTTGAAAATACGATGTTTT 591

Qy    776 TGTGCCAAATAATCCACCATGCCATCGTGCCGATCAAGATGACGAGAGACGATTTAAA 835
Db    590 GCTTGGGTGATACCCCAGCAATATTTCCATGACGGGTCTAAAAATGGTTCTTCACTCATATTGAGA 531

Qy    836 ATTGTTCGCCCCTTGAGCTGCCAAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCA 895
Db    530 ACCACTGCTACCTCCAGCCGTTTAAGTGAGAGGTGGTTCTAACTTGAAAGCTTGATAACTCA 471

Qy    896 AGGACCACCAAATCCATATATTTGGATCGTTAATAATTCAAGTCGCAGGGGTGCCAATATTA 955
Db    470 AGGACTAGTAGCTC---TGGGTGATCTTTAGTAAATCTTAACGCTGGTTGCCTGCAAGATTG 414

Qy    956 CCGCCGGACCAACAATCCATGCCCTGCACATGTTTTGCCCATCTCGCTTACTAATGTCGTAACA 1015
Db    413 CCGCCTACTGCCAAACTTTCTTACCTTGCGATCCTTAGCCCATTAAAAGTCGGTACA 354

Qy   1016 GTGGCTTTGGGCAATTTGAGCCGTGATGGCGATGGAGTGGGTCGTGTTGTTGTTGCTTGAGGTAGA 1075
Db    353 GTACTTTTTTGCATTGGAACCTGTAATCGCCACAATCGGCACAATCAGTGGC------ 304

Qy   1076 TCTCGTGCTTGAGTGTATCANTGAACANAATTGAAACATCGCTGATCACAGGAATGCCTTGA 1135
Db    303 ------------ACGACGCATAAATTGGATATCACCACACAGAGAAATACCTTTA 261

Qy   1136 GCTTTGGCGGCGGCGACAATACTCCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGA 1195
Db    260 GCAAATAGCTGCCTGAAATTTCCGGTAATTGTGGTGTGCAAGGCGCTTAAAATAAATCTCT 201

Qy   1196 TCGGGCTGAGCTTAAGAGTTCACTGTCTCCAAGCTGCCAAAATGACAAATGAAACCCTGTGGGC 1255
Db    200 TCTGCTTGTAATTAATTATTCTTGATCAAGCTGACCAAAACTCGTTTTAAGCCAGCCGGA 141

Qy   1256 AGTTTATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCCTAACTTTGTGACCGGTGA 1315
Db    140 ATCTGATCGTGTCCGGGAGGTGTGGGGGGAGTCCGGTTCGCAGCGACACTTGGTTGACCTTGT 81

Qy   1316 TTGACCAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCATAT 1372
```

```
Db          80 TCATGCAGGAAATTTACAGCAGAAACACCTGATATTCCCAAGCCTGCTACAAACTTTT 24

RESULT 5
ADA29995/c
ID   ADA29995 standard; DNA; 1416 BP.
XX
AC   ADA29995;
XX
DT   20-NOV-2003 (first entry)
XX
DE   DNA encoding Acinetobacter baumannii protein #1282.
XX
KW   ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
KW   vaccine; plant biocontrol agent.
XX
OS   Acinetobacter baumannii.
XX
FH
PN   US6562958-B1.
XX
PD   13-MAY-2003.
XX
PF   04-JUN-1999;   99US-00328352.
XX
PR   09-JUN-1998;   98US-0088701P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Breton G, Bush D;
XX
DR   WPI; 2003-576092/54.
DR   P-PSDB; ADA34121.
XX
PT   New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT   for diagnosing a bacterial disease, as components of antibacterial
PT   vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT   plants.
XX
PS   Example; SEQ ID NO 1282; 328pp; English.
XX
CC   The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC   The A. baumannii nucleic acids and polypeptides are useful as reagents
CC   for diagnosing a bacterial disease, as components of antibacterial
CC   vaccines, as targets for antibacterial drugs, to detect the presence of
CC   A. baumannii and other Acinetobacter species in a sample, in screening
CC   compounds for the ability to interfere with the A. baumannii life cycle
CC   or to inhibit A. baumannii infection, and as biocontrol agents for
CC   plants. The present sequence represents DNA encoding an A. baumannii
CC   protein.
XX
SQ   Sequence 1416 BP; 417 A; 262 C; 326 G; 411 T; 0 U; 0 Other;

Query Match             16.4%;  Score 226.2; DB 9;  Length 1416;
Best Local Similarity   52.1%;
Matches  718;  Conservative    0;  Mismatches 593;  Indels  66;  Gaps   7;
```

(alignment follows)

```
Qy          2 AAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATCC  61
              ||||  |       || |   | | || |||  |||  || |    || |  ||||| |||
Db       1409 AACGAATTCACGCAGCAGGCAACAAACTGCTGACCCGGTCATTCATTAAACTTTTAAACATATCA 1350
```

```
Db    149  TCATGCAGGAAATTTACAGCAGAAAACACCTGATTATTCCCAAGCCTGCTACAACTTTT 93

Db    RESULT 6
XX
AAS53453/c
ID    AAS53453 standard; DNA; 1314 BP.
XX
AC    AAS53453;
XX
DT    13-FEB-2002 (first entry)
XX
DE    Haemophilus influenzae DNA for cellular proliferation protein #235.
XX
KW    Antisense; ds; prokaryotic cellular proliferation gene; antibiotic;
KW    antibacterial; drug design.
XX
OS    Haemophilus influenzae.
XX
PN    WO200170955-A2.
XX
PD    27-SEP-2001.
XX
PF    21-MAR-2001; 2001WO-US009180.
XX
PR    21-MAR-2000; 200US-0191078P.
PR    23-MAY-2000; 200US-0206848P.
PR    26-MAY-2000; 200US-0207727P.
PR    27-OCT-2000; 200US-0245578P.
PR    27-NOV-2000; 200US-02536250P.
PR    22-DEC-2000; 200US-0257931P.
PR    16-FEB-2001; 2001US-0269308P.
XX
PA    (ELIT-) ELITRA PHARM INC.
XX
PI    Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
PI    Yamamoto RT, Xu HH;
XX
PR    WPI; 2001-611495/70.
DR    P-PSDB; AAU35594.
XX
PT    New polynucleotides for the identification and development of
PT    antibiotics, comprise sequences of antisense nucleic acids.
XX
PS    Claim 27; SEQ ID NO 7090; 51lpp; English.

CC    The invention relates to antisense inhibitors of genes essential to
CC    prokaryotic cellular proliferation, their use in identifying the genes,
CC    their use in the discovery of novel antibiotics, the essential genes
CC    themselves and the encoded proteins. The prokaryotes used are Escherichia
CC    coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae,
CC    Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also
CC    useful for the identification of potential new targets for antibiotic
CC    development. The antisense nucleic acids can also be used to identify
CC    proteins used in proliferation, to express these proteins, and to obtain
CC    antibodies capable of binding to the expressed proteins. The proteins can
CC    be used to screen compounds in rational drug discovery programmes. The
CC    antisense nucleic acid sequence is also useful to screen for homologous
CC    nucleic acids which are required for cell proliferation in a wide variety
CC    of organisms. The present sequence encodes an essential prokaryotic
CC    cellular proliferation protein. Note: The sequence data for this patent
CC    did not form part of the printed specification, but was obtained in
CC    electronic format directly from WIPO at
CC    ftp.wipo.int/pub/published_pct_sequences
XX
SQ    Sequence 1314 BP; 403 A; 214 C; 294 G; 403 T; 0 U; 0 Other;

Query Match    11.6%;  Score 160.4;  DB 4;  Length 1314;
Best Local Similarity  49.6%;  Pred. No. 2.4e-35;
Matches  683;  Conservative  0;  Mismatches  606;  Indels  89;  Gaps  7;

Qy      1  TAAAGCGTGAACCAYGTTAACAAAACTTATCACCTGCCTCAHTATTAATTTTTAAATTGATC  60
           ||||    |||  |    | |||||  ||| | ||||||||  ||| |  | | || ||||||
Db   1308  TAAACATTGAGCTAAATGCTAAAATTCTTCGCCGCGGCTTTTCAAAAAGAAGACAAACTGATC  1249

Qy     61  CATGCTAGCCACAAGCACAGGTGACAAAAGGCACAGCCCTAGCGGTGCTTTTGACTGGCTGCCAA  120
           ||||||      ||||| |||||  ||  | |  | ||     ||  |  | ||||||| ||||||
Db   1248  GAGACTTTGCACAAGCAGGCGGACAAATAATACCA------------TATCTCCGCTTTTTGCAA  1201

Qy    121  TACTGCCCGATCAAGTGTTACCAACCTGATATATTTTAGATGCATCACCTTTT  180
           |||||  ||||  || |  ||  ||||||| |||||||| |||||| |  |||
Db   1200  TGTTGGGGCGTAAAAAATTCTATCGCTTGTTCCATTGTATCGAACAA------ATAACTTTG  1147

Qy    181  GAGTAAATCTTCGCTCAHTATTATGGGGCATCTTGACCGATTAAATACACATCCACCCA  240
           | || ||||| ||| ||| | |||||||||| |||||  ||||||  ||||| ||||||
Db   1146  CGATGAAAATTTTTGCAGCAGCAGGCACCATCTCGACCAAACAATAACAAATAATGTGTGG  1087

Qy    241  CTTGCATAACAATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCAAAAT  300
           ||||| |||||| | ||||| |||| || |||||||||||| ||| ||||| || ||||||
Db   1086  TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTCCGGTCTCGCCGCTAGCAA  1027

Qy    301  CAATGCCAGTGATGATGATAATGCCATAAAACCGTGCCAAGCCCATCAATTGCTGCAATGGT  360
           ||||| | | | |  |||||| ||||  ||||||||||||||||| |||  | ||||||||
Db   1026  CAAATGCAAATTTACCCTTCAAATATA-------AAGCCCAGCCCAAATGCAGCAAACTGTACT  976

Qy    361  GCCGACATGTTTGTTGTACCTTTTTGAATCATTAAAAATATGCCTTGCTGTCATCAATATCAATATA  420
           |||||||||||| |  || ||||||||| ||||  |||  |||| ||||| ||||||||||||||||
Db    975  CCCCACATTTGTTTGCTTTGAGTCATTAATCCAACGAAATGCCATTAGCTTGAGCACTAA  916

Qy    421  TTCACACGGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG  480
           |||||| |  ||| | ||||| |||||||||||| || || |||| || ||||| |||||
Db    915  TTGAAAACGAATGAATCTAAACCCTTTGAAAATGACGAAGTGCGGTACGAATTGAAATCTAAATT  856

Qy    481  AAGATTCGGCAGCTGTGCCTAAGGCAAGTGCACAGATAAGGCAATTGAGTAGGGTTATGCTTGCC  540
           ||||||  |||| |||||    ||||   | |||  |||||||   |||| |||  || | ||||
Db    855  AATACCTTATAGCTTGGTCGTCGCCAATTGCTGGTGCGTTCATCTTAATAATTATTGCGGACC  796

Qy    541  TTTGATTTTTAATTGATGGATTGGTTGGGTAAAAGTTTTTGTTTTTTTTGATGATATAATGCCATACC  600
           |||    |||    | |  | |||||||||| | |||||||||    |||||||| ||  |||||||
Db    600  AACCAATGTAGCTTCTTCACAAGGTAAAAATCACTTTCATTCTTTTTACCCATTAAATATATGCGTT  736

Qy    601  ATCACCATCATCAAATAAAATAAAATCTGC--CGTTTGGTGTGGGCTAAGTAAGCTAGTTATTGTTG  659
            |||||||||| ||| |||    | |||||   | |||  ||| ||| |||| ||| || ||||| |
Db    659  GCCATTTTCAGTTTTTTAGCGACTAATAACTCGCACTATTTTCCGCAAAAGAAAACGGTATGTTT  676

Qy    660  TGATAATGGGTGTGATTTGGGGCGACCGCTTGTCTGCAGTGTGGGTCAAGCATTGCTTGGGCTAAAT  719
           |||         ||||| |||||      ||||||||| || | ||||||||| ||||||| ||||
Db    735  GCCATTTCTTCAGTTTTTTAGCGAATAATCCGCACTATTTTCCGCAAAAGAAGAAAACGGTATGTTT  676

Qy    675  CGCTTGATTTTCGTTTTCCCCAAAAGTCAGCC-------  644
           ||||                  ||||| ||||
Db    735  TGATAATGGGTGTGATTTGGGGCGACCGCTTGTCTGC  719

Qy    720  TGGCACTCGTCTTGGTAGATGATGGGCGGTTTGGCAATTTCAAAAATACGGCAATTTTTTTGTG  779
           ||||||  |||  | |||| ||  ||||   ||||||   ||| ||| || |||||||  ||||
Db    643  ---TATCTTCATTGTTCAACACACTACTTTAGCATTATTATGATATTTTAATTTTCGTT  587

Qy    780  CCAAAATAAATCCACCATGCCAGCCATCGTGCCGATCAAAGATGATCAGCGGAGAGAGATTTAAAATTG  839
           |||||||   || |||| || | ||||||||| || |||||||| |||||  || |||||  |||||
Db    586  GGCGATAATCTTCTAAATCCATATAGCGGATCCATATATGATGATCTTCAGTCGACAGTTCAAGACAG  527

Qy    840  TTGCCCCTTGAGCTGCCAAAAATTTGAGATATGCCTCAAGCTGAAAAACTTGACAGCTCAAGCA  899
           ||      || ||| | |||  |||||||| | ||  |||||  |||||||| || || |||||
Db    526  TCGCTCGCGCAGCCTTTTAAGCTTATAAGTTGTCTCAAGCTGAAAACGCTAGTA  467

Qy    900  CCACCAAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGGGTGCCAAATTATTACCGC  959
           ||||| |||  ||| || | ||| | ||| ||| ||| | | |||| | |||| |||||||
Db    466  CATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGAACCCCAATAATTTCCGC  410

Qy    960  CGACACCCAACAATCATGCCTGCAGCCAGCTTTCGCCATTCAGTTTCGCCTACTAAGTCGTAACAGTGC  1019
           ||||||| |||| | || ||  | |||||||| | |||  | ||| ||| ||| || |||||||||
Db    409  CCATACCAACTTTCACACCAGCAGCTTTCGCCATTTCATAAAACTAAAGTAGTTACGGTAC  350

Qy   1020  TTTTGGCATTTGAGCCTGTCGATGGCCGATGACTGGTGTGTTGCTTTGACTTAGATCTC  1079
           ||||    |||||||  |||| ||  |  | | |||| | |||| ||||| ||| ||||
Db    349  TTTTACCATTTGAACCTGTAATCCCACCAACAATTGGCTTTGTCGCCTGGGCGGGCA------  298

Qy   1080  GTGCTTTGAGTGTATCAAATGAACAATTGAACATTCGCTGATCACAGGGAATGCCTTGAGCTT  1139
                  |||| ||| ||||||| ||||| |||||| || || |||| ||||| | || |||
Db    297  --------GAATAATTCAATATCGGCGGATTAGTACTTCCACTCCGGCTTTAA  257
```

```
Qy         1140 TGGCGGGACGAATACTCGGCGGCGGTTCTTGGGTTAATCCGGGGGCCTGATGATCTCGATCGG 1199
                |||| |||||||   ||| |||| | |||||| |||| |||||| || ||| ||||||||||
Db          256 GTGCGGTTTGAAATTCTCGTGTGTTTTACCGGCCAGCCCTGGCGTAATAACAATCATATCGC 197

Qy         1200 CTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAAATGACANATGACGACCGTGGGCAGTT 1259
                |||  |||  | ||   |||  || ||| ||||  ||| || ||||| ||   |||||||||
Db          196 TTTCAAGTAACCAATTCCTGATTTAAACTACCAGTATGAAGGGGATAATTTTGAGGAAGTT 137

Qy         1260 TATCTGCCAAGGTAGGAATTGGGGATTAACCATTCGATGAGATGCTAACTTTGTGACCGTGGA 1319
                |||| | |||| |||| ||   ||||| ||||| ||| | | |||| ||||| |||| |||
Db          136 TATCAATACCAGTAGATTTTTTCGAGTATCAATCACACACGAATATTAGCCTGTTGGGATA 77

Qy         1320 CCAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCAATATTTTTT 1377
                |||||| |||||| | |||| |||| |||  || || | ||||  |||||||||||||
Db           76 AGAGATAATCCACACAAGAAAGACCTGTTTTGCCAAGCCCGATGATGATATATTTTT 19
```

RESULT 7

ACA34366/c	
ID	ACA34366 standard; DNA; 1314 BP.
XX	
AC	ACA34366;
XX	
DT	19-JUN-2003 (first entry)
XX	
DE	Prokaryotic essential gene #16023.
XX	
KW	Antisense; ds; prokaryotic essential gene; cell proliferation;
KW	drug design; gene.
XX	
OS	Haemophilus influenzae.
XX	
PN	WO2002277183-A2.
XX	
PD	03-OCT-2002.
XX	
PF	21-MAR-2002; 2002WO-US009107.
XX	
PR	21-MAR-2001; 2001US-00815242.
PR	06-SEP-2001; 2001US-00948993.
PR	25-OCT-2001; 2001US-0342923P.
PR	08-FEB-2002; 2002US-00072851.
PR	06-MAR-2002; 2002US-0362699P.
XX	
PT	New antisense nucleic acids, useful for identifying proteins or screening
PT	for homologous nucleic acids required for cellular proliferation to
PT	isolate candidate molecules for rational drug discovery programs.
XX	
PA	(ELIT-) ELITRA PHARM INC.
PI	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX	
DR	WPI; 2003-029926/02.
DR	P-PSDB; ABU30496.
XX	
PS	Claim 14; SEQ ID NO 22236; 1766pp; English.
XX	
CC	The invention relates to an isolated nucleic acid comprising any one of
CC	the 6213 antisense sequences given in the specification where expression
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:
CC	(1) a vector comprising a promoter operably linked to the nucleic acid
CC	encoding a polypeptide whose expression is inhibited by the antisense
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC	polypeptide or its fragment whose expression is inhibited by the
CC	antisense nucleic acid; (4) an antibody capable of specifically binding
CC	the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC	proliferation or the activity of a gene in an operon required for
CC	proliferation; (7) identifying a compound that influences the activity of
CC	the gene product or that has an activity against a biological pathway
CC	required for proliferation, or that inhibits cellular proliferation; (8)
CC	identifying a gene product required for cellular proliferation or the biological
CC	pathway in which a proliferation-required gene or its gene product lies
CC	or a gene on which the test compound that inhibits proliferation of an
CC	organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC	compound's activity; (11) a culture comprising strains in which the gene
CC	product is overexpressed or underexpressed; (12) determining the extent
CC	to which each of the strains is present in a culture or collection of
CC	strains; or (13) identifying the target of a compound that inhibits the
CC	proliferation of an organism. The antisense nucleic acids are useful for
CC	identifying proteins or screening for homologous nucleic acids required
CC	for cellular proliferation to isolate candidate molecules for rational
CC	drug discovery programs, or for screening homologous nucleic acids
CC	required for proliferation in cells other than S. aureus, S. typhimurium,
CC	K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC	prokaryotic essential genes. Note: The sequence data for this patent did
CC	not form part of the printed specification, but was obtained in
CC	electronic format directly from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences
XX	
SQ	Sequence 1314 BP; 214 A; 294 G; 403 T; 0 U; 0 Other;
```

```
Query Match             49.6%;  Score 160.4;  DB 8;  Length 1314;
Best Local Similarity   49.6%;  Pred. No. 2.4e-35;
Matches   683;  Conservative    0;  Mismatches  606;  Indels   89;  Gaps    7;
```

```
Qy            1 TAAAGGCTAGCACAAGCATGTTAACAAACTTATCACCTCGCTCATTATATAATTTTAAATTGATC 60
                |||| || ||| ||  |||| || || |  || ||| || || | || ||  | ||| ||||||
Db         1308 TAAACATTGGAGCTAAATGCGTAAAAATTCTTCGCGCGGCTTTTCAAAGAGAGCAAACTGATC 1249

Qy           61 CATGCTTAGCACAAGCAGGGTGACAAAAGCACAGCCCTAGCCGGTGCTTTGACTGGCTGCCAA 120
                |||||||  || ||| || ||| ||   ||| ||||| || || |||| || ||| ||  |||
Db         1248 GAGACTTGCGACAAGCAGGCGGCGGCACAACAATTAAATACCA------TATCTCCGCTTTGCAA 1201

Qy          121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAATGTGCATCACCTTT 180
                || ||||| || |||  ||| || || ||||||||  ||| || |||| |||| || ||||
Db         1200 TGTTGGGGTGTAAAAAATTCATCGCTTGTTCCATTGTATCGGAACAA------ATAACTTTG 1147

Qy          181 GAGTAAAATCTTGCTCCAATTATCCGGGGGCCATCTTGACGCGGATTAAATACACACTATCCACCCA 240
                ||| ||  |||| ||| ||||| || ||| ||  || || || ||| ||| |||| | ||||| |
Db         1146 CGATGAAAAATTTTGCAAGCAGCGGCACCCATTCTGACCAAAAACAATAACAAATAATGTGTGG 1087

Qy          241 CTGCATAACATATGCGTGACAGCTCACTAAAAATCTTGACCTTTTACCAAGCCACCCAAAAT 300
                || || |||||||||||||  || | || |||| ||| ||||| || ||||| ||||||||
Db         1086 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCCGGTCTCCCGCCTAGCAA 1027

Qy          301 CAATGCCCAGTGATGATTATGCCATAAAACCGTGCCAAGCCCATCAATTGCTGACAATGGTTGA 360
                || ||| || || || |||| ||| | |||||   | |||||| |||||| |||||| |||
Db         1026 CAAATGCAATTACCCTCAATATA--------AAGCCCAGCCAATGCAGCAACTGTACT 976

Qy          361 GCCGATGTTTGGTACCTTTTGAATCATTAAAATATGGCCTTCTACATCACAATATA 420
                ||| ||||| || ||||||  |||| ||| ||| |||| || || |||| |||| |
Db          975 CCCCACATTTGTTGCTTTAGAGTCATTAATCCCACGAAATGCCCATTAGCTTGATGCACTAA 916

Qy          421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCGCATCCATGGG 480
                ||| || ||| || ||| ||| ||||| || ||||| || |||||| || ||||||
Db          915 TTGAAAACGATGATCTAACCCTTTGAAATGCGGTACGAAGTGCGCAAGAATCTAAATT 856

Qy          481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCAGCATTGAGTAGGTTATGCTTGCC 540
                || |||| ||||| || ||||| ||||  |||| |||||| ||||| |  || || |||||
Db          855 AATACCTATAGCTTCTGCGCAAATGGTTGCTGCCAAAATGTTCATTATAAATTATGCGACC 796

Qy          541 TTTGATTTTAATTGATGGANTTGCTGGTGTAAGAAGTTGTTTTGTATTGGGAATGGCCATACC 600
                |||||||| || |||| || |||||  || ||||| |||||| ||| ||| ||||||||||
Db          795 AACCAAATGTAGCTCTTCAAGGTAAAAATCACTCACTTTTACCATTAAATTAATTAAGTTGCTT 736

Qy          601 ATCACCATCATCAAATAAATAAAAATAAAAACTTGC--CGTTTGGTGGCTAAGCTATTGTTG 659
                |||| ||   |||||| || || |||| ||| ||  ||| ||||  || |||||| || |||
Db          735 GCCAATTTCAGTTTTAGCCAATAATCCGCACATTTTCCGCAAAAGAAACGGTATGTTTT 676

Qy          660 TGATAATGGTTGGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAGCATTGTGGCTAAAT 719
                |||| ||   |||| |||||  |||  ||||| ||| || || ||||| || || |||
Db          675 CGCTTGATTTTCGTTTTTCCCCAAAAAGTCAGCC---------------644

Qy          720 TGGCATCGTCGTTCGTTGGATGATGAGCGGGGGTTTGGCAATTTTCAAAAATACGGCAATTTTGTG 779
                ||| ||  || || |||| |       |  | ||  || |||||| ||   || | | |||
```

```
Qy    643 ---TATCTTCATTGTTCAACACCTACTTTAGCATTATGATATAAAATGCGTAATTTTGCTT  587
             ||||||||||||||||||||
Db    780 CCAAATAATCCACATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTAAAATTG  839

Db    586 GGCGATAATCTTCTAAAATCCATATAGCGATCCATATATGATCTTCAGTCACGTTCAAGACAG  527

Qy    840 TTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGGTCAAAAACTTGACAGCTCAAGCA  899
Db    526 TCGCTCCGCGCAGCGTTTTAAGCTATATAAGCTGTCTCAAGCTGAAAAACTAGAAAAGCTCTAGTA  467

Qy    900 CCACCAAATCCATATATTTTGGATCGTTAATAATTCAAGTGCAGGCGGTGCCAAATTATTACCGC  959
Db    466 CATAAAGTTCACAAACTTCATTC---AACAATGACAAAGCGGGAACCCCAAATATTTCCGC  410

Qy    960 CGACACCAACAAATCATGCCGCATGTTTTGCCCTACTACTAATGTCGTAACAGTGC  1019
Db    409 CCATACCAACWTTCACCAGACCAGCTTTCGCCATTTCATAAACTAAAGTAGTTACGGTAC  350

Qy   1020 TTTTGGCATTTGAGCCTGTGATGGGCGGATGGTTGGTTGCTTTTGTTGTTTGTTGAGTTAGATCTC  1079
Db    349 TTTTACCATTTGAACCTGTAATCCCCACAATTGGCTTTGTGCTGGGCGGCA---------  298

Qy   1080 GTGGCTTTGAGTGTATCAAATGAACAATTGACACATTCGCTGATCACAGGAAATGCCTTGAGCTT  1139
Db    297 --------GAATAAATTCAATATATCGCCGATTACTTCCACTCCCGCTTTAA  257

Qy   1140 TGGGCGGCGACAAATACTCGGCCGTTCTTCGGGGGTTAATGCCGGGCGGTCGATGATCTGATTCGG  1199
Db    256 GTGGCGGTTTGAGATTTCTGGTGTTTTTACCGGAAGCCCTGGTACTAATAACAATCATATCGC  197

Qy   1200 CTGAGCTTAAGAGTTCACTCTGCTGCCAAAAATGACAAATGAACGGCCTGTGGGCCAGTT  1259
Db    136 TATCAATACCAGTAGGATTTTTCGAGTATCAATCACACGAATATTAGCCTGTTGGGATA  77

Qy   1260 TATCTGCCAAGGTAGGATTGGGATTACCATCGATGANGATGCTAACTTTGTTGACCGTGATTGA  1319
Db    196 TTTCAAGTAACCATTCCTGATTTAAACTACCACGTATGCAAGAGGGATATTTTGAGGAAGTT  137

Qy   1320 CCAAAAAATTCACAGCTGACAGTCCAAAACCGACGACGAGGGCATAATTTTTTT  1377
Db     76 AGAGTAATCCACACAAGAAAGAACCTGTTTTGCCAAGCCCGATGATAGTAATATATTTTT  19
```

```
Query Match    11.6%;  Score 160.4;  DB 11;  Length 1314;
Best Local Similarity   49.6%;  Pred. No. 2.4e-35;
Matches  683;  Conservative  0;  Mismatches 606;  Indels 89;  Gaps  7;

Qy      1 TAAAGCGTGAACCATWGTTAACAAACTTACCACCTCGCTCATTATAATTTTTAAATTGATC  60
Db   1308 TAAACATTGAGCTAAATGCGTAAAAATTCTTCGCCGGGCTTTTCAAAAGAGAAGAACTGATC  1249

Qy     61 CATGCTTAGCACAAGCAGGGTGACAAAAGCACAGCCCTTAGCGCGTGCTTTGACTGTGCCA  120
Db   1248 GAGACTTGCACAAGCAGGCGACACAATAATACCA-----------TATCTCCGCCTTGCAA  1201

Qy    121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCGAACCAAAATGTGGCATCACCTTT  180
Db   1200 TGTTGGGCGTAAAAAATTCTATCGCTTGTTCCATTGTTATCGAACCA------ATAACTTTG  1147

Qy    181 GAGTAAATCTTGCTCAATTATCGGGGCCATCTTGACCGGATTAAAATACACTATCCACCCA  240
```

```
Db   1146 CGATGAAAATTTTGCAAGGAGCGAGCACCATCTCGACCAAAACAATAACAAATAATGTGTGG
          |||               |||                      |||       |||
Qy    241 CTGCATAAACATATGCTGACAGCTCACTACAAAATCTTGACTTTACCAAGCCCACCCAAAAT 300
Db   1086 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTCCGTCTCCGCCTAGCAA 1027

Qy    301 CAATGCCAGTGATGATATGGCCATAAACCGTTGCCAAGCCCATCACAATTGCTGCAATGGTTGA
Db   1026 CAAATGCAATTTACCCTCAATATA----------AAGCCCAGCCAATGCAGCCAACTGTACT 976

Qy    361 GCCGATGTTTGTGTTACCTTTGGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
Db    975 CCCCACATTTGTTGCTTTAGAGTCATTAATCCACGAAATGCCATTAGCTTGATGCACTAA 916

Qy    421 TTCACAGCGATGCGGTTAGCCCCTTTGAAAAGTCTTAAGGGGTATCAAGCCATCCATGGG 480
Db    915 TTGAAAACGATGATCTAACCCTTTGAAATGACGGAAGTGCGGTACGGAATTGAATCTAAATT 856

Qy    481 AAGATTGGCAGCTGTGGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATCTGCTTGCC 540
Db    855 AATACCTATAGCTTGTGCCAAGTCGTGTTGCTGCCAAAATGTTCATATATAATTAATGCGACC 796

Qy    541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTTGTTTTTTTTGATGATATAAATGCCATACC 600
Db    795 AACCAATGTAGCTTCTTCACAAGGTAAAATCACTTCATCTTTTTACCATTAAATATNTGCTTT 736

Qy    601 ATCACCATCATCAAATAAAATAAAAATCTGC-CGTTTGGTGGCTAAGTAAGCTATTTGTTG 659
Db    735 GCCAATTTCAGTTTTAGCCAATAATCCCGACTATATTTCCGACAAAAGAAACGGTTATGTTT 676

Qy    660 TGATAATGGTGTGANTTGGGGCAACGCTTGTCTCAGTGTGGTCAAGCATTGCTTGGCTAAAT 719
Db    675 CGCTTGATTTTGCGTTTTCCCAAAAAGTCAGCC---------------------------- 644

Qy    720 TGGGCATGCTGCTTGGTAGAATGGCGGGTTTGGGCAATTTCAAAAAATACGGCAATTTTTGTG 779
Db    643 ---TATCTTCATTCGTTCACACACTACTTTAGATAATGGGCTAATTTTTTGCTT 587

Qy    780 CCAAAATAATCACCATGCCATGCGTTGCCGGATCAAGATGATCAGCCGGAGAGATTTAAAATG 839
Db    586 GGGGATAATCTTCTAAAATCCATATAGGCGGATTCCGCAAAAGAAAACGGTTATTGTT 527

Qy    840 TTGCCCCTTGAGCTGCCAAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA 899
Db    526 TCGCTGCCCGCCAGCTTTAAGCTATAAAGTGTCTCAAGCTGAAAAACTAGAAAAGCTTCTAGTA 467

Qy    900 CCACCAAATCCATATTTTGGATCGTTAATAATGCCAGGCGTGCCAATAATTACCGC 959
Db    466 CATAAAAGTTCACAAATCTTCATTC---AACAATGACAAACGGGAATCCCAATATTTCCG 410

Qy    960 CGACACCAACAAATCATGCCTGCATGTTTGCCATCTCGGCTACTAATGTCGTAACAGTGC 1019
Db    409 CCATACCAACTTTCACCAGCAGCTTTCGCCATTTCTATCAAGGTAGTAGTTACGGTAC 350

Qy   1020 TTTTGGCATTTGAGCCTGCATGGCGGCGATCATTGGTGTGTTCTTGAGTTAGATCTC 1079
Db    349 TTTTACCATTTGAACCTGTAATCCCCAGTTAAACTAACTGCTTTGCGGGCA 298

Qy   1080 GTGCTTTGAGTGTATCAATGAACANTTGAACACTCGGATTACTTCCACTCCCGCTTTA 1139
Db    297 ----------------------GAATAATTCAATATGCCGATTACTTCCACTCCCGCTTTA 257

Qy   1140 TGGCGGGCGACAACTACTCGGCGTTCTTGGGGTTAATGCGGGGCTGATGATGAATCATCGG 1199
Db    256 GTCGCGGTTTGAATTTCTGGTGTTTTTTACGGCAACCCTCGGGCTAATAACAATCATATCGC 197

Qy   1200 CTGAGCTTAAGAGAGTTTCACTCGTCCAAAAATGACAAATGAAACGCCTGTGGGCAGTT 1259
Db    196 TTTCAAGTTAACCATTCCTGATTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 137

Qy   1260 TATCTGCCAAAGGTAGGATTGGGATTACCATCGATGAATGCTAACTTGTGACCGTGATTGA
          |||||
Db    136 TATCAATACCAGTAGGATTTTTTCGAGTATCAATCATCACCAGCGAATATTAGCCTGTTGGGATA 77

Qy   1320 CCAAAAAATTCACAGCTGACAGTTCCAGAGCTGTGCCCAAACCGACGACGGGCAATATTTTTT 1377
Db     76 AGAGATAATCCACACAAGAAAAGACCTGTTTTTGCAAGCCCGGATGATTAGTAATATATTTTT 19
```

RESULT 9
AEC10868/c

```
ID   AEC10868 standard; DNA; 1314 BP.
XX
AC   AEC10868;
XX
DT   20-OCT-2005 (first entry)
XX
DE   Haemophilus influenzae MURD DNA.
XX
KW   protein purification; antibacterial; antimicrobial; infection;
KW   drug screening; gene; d; UDP-N-acetylmuramoylalanine-D-glutamate ligase.
XX
OS   Haemophilus influenzae.
XX
PN   US2005181388-A1.
XX
PD   18-AUG-2005.
XX
PF   04-OCT-2004; 2004US-00958216.
XX
PR   02-APR-2002; 2002US-0369511P.
PR   04-APR-2002; 2002US-0369817P.
PR   04-APR-2002; 2002US-0370102P.
PR   08-APR-2002; 2002US-0370778P.
PR   08-APR-2002; 2002US-0370792P.
PR   08-APR-2002; 2002US-0370820P.
PR   08-APR-2002; 2002US-0370859P.
PR   08-APR-2002; 2002US-0370999P.
PR   08-APR-2002; 2002US-0370915P.
PR   09-APR-2002; 2002US-0371067P.
PR   09-APR-2002; 2002US-0371107P.
PR   09-APR-2002; 2002US-0371140P.
PR   09-APR-2002; 2002US-0371185P.
PR   31-MAY-2002; 2002US-0385426P.
PR   04-JUN-2002; 2002US-0385751P.
PR   05-JUN-2002; 2002US-0386018P.
PR   05-JUN-2002; 2002US-0386367P.
PR   05-JUN-2002; 2002US-0385548P.
PR   05-JUN-2002; 2002US-0385553P.
PR   05-JUN-2002; 2002US-0385566P.
PR   05-JUN-2002; 2002US-0385777P.
PR   06-JUN-2002; 2002US-0386283P.
PR   06-JUN-2002; 2002US-0386390P.
PR   06-JUN-2002; 2002US-0386430P.
PR   06-JUN-2002; 2002US-0386601P.
PR   06-JUN-2002; 2002US-0386826P.
PR   06-JUN-2002; 2002US-0368699P.
PR   31-JUL-2002; 2002US-0399972P.
PR   01-AUG-2002; 2002US-0400348P.
PR   05-NOV-2002; 2002US-0424053P.
PR   06-NOV-2002; 2002US-0424380P.
PR   06-NOV-2002; 2002US-0424395P.
PR   08-NOV-2002; 2002US-0425086P.
PR   08-NOV-2002; 2002US-0425200P.
PR   24-DEC-2002; 2002US-0436288P.
PR   24-DEC-2002; 2002US-0436345P.
PR   24-DEC-2002; 2002US-0436349P.
PR   26-DEC-2002; 2002US-0436560P.
PR   26-DEC-2002; 2002US-0436567P.
PR   26-DEC-2002; 2002US-0436568P.
PR   27-DEC-2002; 2002US-0436675P.
PR   27-DEC-2002; 2002US-0436708P.
PR   27-DEC-2002; 2002US-0436734P.
```

```
PR   27-DEC-2002; 2002US-0436804P.
PR   27-DEC-2002; 2002US-0436834P.
PR   27-DEC-2002; 2002US-0436842P.
PR   27-DEC-2002; 2002US-0436861P.
PR   27-DEC-2002; 2002US-0436885P.
PR   27-DEC-2002; 2002US-0436889P.
PR   27-DEC-2002; 2002US-0436893P.
PR   27-DEC-2002; 2002US-0436900P.
PR   27-DEC-2002; 2002US-0436947P.
PR   30-DEC-2002; 2002US-0436971P.
PR   30-DEC-2002; 2002US-0436987P.
PR   30-DEC-2002; 2002US-0437013P.
PR   30-DEC-2002; 2002US-0437038P.
PR   31-DEC-2002; 2002US-0437141P.
PR   31-DEC-2002; 2002US-0437281P.
PR   31-DEC-2002; 2002US-0437527P.
PR   30-DEC-2002; 2002US-0437620P.
PR   31-DEC-2002; 2002US-0437638P.
PR   02-APR-2003; 2003WO-CA000462.
PR   04-APR-2003; 2003WO-CA000464.
PR   08-APR-2003; 2003WO-CA000481.
PR   08-APR-2003; 2003WO-CA000485.
XX
PA   (AFFI-) AFFINIUM PHARM INC.
XX
PI   Edwards A, Dharamsi A, Vedadi M, Alam MZ, Arrowsmith C, Awrey DE;
PI   Beattie B, Buzadzia K, Canadien V, Domagala M, Houston S;
PI   Kanagarajah D, Li Q, Mansoury K, Mcdonald M, Nethery-Brokx K, Ng I;
PI   Ouyang H, Pinder B, Richards D, Tai M, Thalakada R, Vallee P;
PI   Virag C;
XX
DR   WPI; 2005-628189/64.
DR   P-PSDB; AEC10869.
XX
PS   Claim 65; SEQ ID NO 400; 667pp; English.
XX
PT   New composition comprising purified polypeptides from bacteria (e.g.
PT   Escherichia coli), useful for diagnosing, preventing or treating
PT   microbial infections, or in pharmacogenomic or drug screening procedures.
XX
CC   The invention relates to a composition (I) comprising purified
CC   polypeptides from bacteria. Also described: (1) a crystallized,
CC   recombinant polypeptide comprising an amino acid sequence of (I), where
CC   the polypeptide is in crystal form; (2) a crystallized complex comprising
CC   the crystallized, recombinant polypeptide and a co-factor or a small
CC   organic molecule, where the complex is in crystal form; and (3) a host
CC   cell comprising a nucleic acid encoding a polypeptide of (I), where a
CC   culture of the host cell produces at least about 1 mg of the polypeptide
CC   per liter of culture and the polypeptide is at least about one-third
CC   soluble as measured by gel electrophoresis. The composition and methods
CC   are useful for diagnosing, preventing or treating diseases, such as
CC   microbial infections. These may also be used in pharmacogenomic or drug
CC   screening procedures. The present sequence represents a Haemophilus
CC   influenzae UDP-N-acetylmuramoylalanine-D-glutamate ligase gene sequence,
CC   which is used in an example from the present invention.
XX
SQ   Sequence 1314 BP; 403 A; 214 C; 294 G; 403 T; 0 U; 0 Other;

Query Match             11.6%;  Score 160.4;  DB 14;  Length 1314;
Best Local Similarity   49.6%;  Pred. No. 2.4e-35;
Matches  683;  Conservative    0;  Mismatches 606;  Indels  89;  Gaps   7;

Qy          1 TAAAGCGTGAAACCATGTTAACAAACTTATCACCTGCTCAATTATAAATTTTAAATTGATC 60
              |||  ||  ||||  |||  |||  |||||||  |||  || ||| |||| |||||||||||
Db       1308 TAAACATTGAGCTAAATGCGTAAATTCTTCGCCCGGCCTTTTCAAAAGAAGCAAAACTGATC 1249

Qy         61 CATCCTAGCACACAGGTTGACAAAAGCACAGCCCTAGCGGTTGCTTGGCTGCCCA 120
              ||||  |  || |   || |  || |||      || || | || ||| |  ||  ||||
Db       1248 GAGACTTGCACAAGCAGGCGCACAATAATACCA---------TATCTCCGCTTTGCAA 1201

Qy        121 TACTGCCCCGATCAAGTGTACCAAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 180
              |||      |||  |||     ||      |||| |||||     ||  ||||   ||||||
Db       1200 TGTTGGGCCGTAAAAATTCTATCGCTTGTTCCATTGTATCGAACAA-----ATAACTTG 1147

Qy        181 GAGTAAAATCTTGCTCAATTATCGGGGGGCATCTTGACGGATTAAATACGACTATCCACCA 240
              ||  ||||| | |  ||||| ||   || ||| || ||  ||||   |  |  ||| ||  |
Db       1146 CGATGAAAATTTTGCAAGGCAGCGCGACCATCTCGACCAAAAACAATAACAAAATAATGTGTG 1087

Qy        241 CTGGCATAACATATGCTGACAAGCTCACTCAAAAATCTTGACCTTTACCAAGCCCACCCAAAT 300
              || |  |||  | | |||||    || ||| ||| |||||      |||  | ||  || ||
Db       1086 TTGATTAATTAATTCAGCTATTCTGAAAAAATCAGCGCCCCTTTTCCGTTCACTTTGCAA 1027

Qy        301 CAAATGGCCAGTGATGATATGCCATAAAACCGGTGCCAAGCGCCATCAAATGTCGAAATGTTGA 360
              ||||  | |  ||  | | ||| |    |  ||  |||||| || |  || || | |
Db       1026 CAAATGCAATTACCCTCAATATA--------AAGCCCAGCCAAATCCAGCAACTGTACT 976

Qy        361 GCCGATGTTGTTGCTTTTTGAATCAATTAAAAATATGCCTTGCTATCAAATATCATCAATATA 420
              ||||| ||  | || ||| |   |||||||||  || | | |||| ||  |  |||  || |
Db        975 CCCCACATTGTTGCTTTAGAGTCATTAAAATCCAACGAAATTGCCATTAGCTTGCACTATGCTA 916

Qy        421 TTCACGACGGATGCCGGTAGCCCCTTTGAAAGTCTTAAGGGTATCAAGACCATGGCCATCCATGGG 480
              ||||  |  ||||||    || |  ||| |   ||| ||   |||   |||  ||| |||   ||
Db        915 TTGAAAACGGATGAATCTAACCCTTTGAAAATGACGAAGGTTACGAAATTGAATCTAAATT 856

Qy        481 AAGATTGGCACGCTGTGCCTAGGCAAGTGCAGATTAAAGGCATTGAGTAGGTTTATGCTTGCC 540
              || |  | |  ||| |   || | || |  |  | ||  |   |||   |||| | ||||
Db        855 AATACCTATAGCTTGTGTCCAAATGCCTGTTGCTGCCAAAAAATGTTCATATATATTATGGCGACC 796

Qy        541 TTTGATTTTAATTGATGGATTGGTAAAAGTTGTGTTTTTTGATGATAATATGCCATACC 600
              ||| |||||||  |  ||||| |  |||||| |  ||| |  |||| ||| | |||  |
Db        795 AACCAAGTAGCTTCTTCACAGTTAAAATCACTTCATCTTTTACCATTAAAATATTGCTT 736

Qy        601 ATCACCATCATCAAATAAATAAAAATCTGC-CGTTTGGTGGCTAAGTAAGCTATTGTTG 659
              || ||||| |  | || | ||   || |  |  | ||||||| | | |  ||| | ||
Db        735 GCCATTTTCAGTTGTTTAGCCAAATAATCCGGACTAATTCCGGCAAAAGAAACGGTATGTTT 676

Qy        660 TGATAAATGGGTGTGATTGTGGGGGCAAGCGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAAT 719
              || |  | | | ||| | | |  |||||| | | ||||||  |||  |||    | | ||
Db        675 CGCTTGATTTCGTTTTCCCACAAAGTCAGCC-------------------------------- 644

Qy        720 TGGCATCGTCTTGGGTAGATGATGGGCGGTTGGCAATTTCAAAAATACGCAAAATTTTTGTG 779
              ||    |  || |  |   |||| |  || | ||  ||||  |||||| || |||  |||||
Db        643 ---TATCTTCATTGTTCACACACTACTTAGCATTCATGATAATAAGCGTAAATTTTGCTT 587

Qy        780 CCAAATAATCCACCATGCCGATCAAGATGATCAGCGGGAGAGGATTTAAAATTG 839
              ||||||||||  | |  ||||| ||  |||  |   |  |   ||||| | ||
Db        586 GGCGATAATCTTCTAAAATCCATATAGCGGATCCATATAGTCACGTTCAAGACAG 527

Qy        840 TTGCCCCTTGAGCGTCGCCCAAAATTTGAGATATGCTCAAGCTCAAAGACTTGACAGCTCAAGCA 899
              ||| | | || | | |  ||| | ||||  |  | || |||| ||  | ||| || |  ||
Db        526 TCGCTCCGCGCAGCTTTTAGCTATAGAAGTTGTCTCAAGCTGAAAAACTAGAAAGGTCTCAGTA 467

Qy        900 CCACCAAATCCATATTTGGATCGTTAATAAATTCAAGTGTGAGGCGTGCCAATATTACCGC 959
              || |  |||  || |  || ||| |  |  || |  | |  |  |||| | || ||| ||
Db        466 CATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAAATATTTCCGC 410

Qy        960 CGACACAGCAAATCCATGCCCTCCATGTTTTGCCATCTCGCCTACTAATGTGTCGTAACAGTGC 1019
              || |  |||  ||||  |  ||| |  |  | || |  | ||  || | ||  |||  | ||
Db        409 CCATACCAAACTTTCACACCAGCAGCTTTCGGCCATTTCATCAAAATAAAGTAGTTACGGTAC 350

Qy       1020 TTTTTGGCATTTGAGCCTTGATGAGCGGATGGCGGATTGGGTGTGTTGTTTGCTTGAGTTAGATCTC 1079
              ||| |  ||| |  |||| |||||||||||  | ||| |  ||| | |  | ||  || | || ||
Db        349 TTTTACCATTTGAACCTGTAAATCCCACAATTGGCTTTGTGCGCTGCGGCA------ 298

Qy       1080 GTGCTTGAGTGTATCAAATGACAATGAACATGCGTGATCGCAGGAAATGCCCTTTAA 1139
              ||| |  |   | |||||   || | | | || | |   || ||||||| |||
Db        297 ------------------GAATAATTCAATATATCGGCGATTCACTCCCGGCTTTGAGCTT 257

Qy       1140 TGGGCGCGACAATACTCGGGCGGGTTTCTTGGGTTAATGCCGGACTGATGATGATGATCGCGG 1199
              ||| |  | ||| |   |  | |  || |   || || |  | | |  || ||| | |||
Db        256 GTGCGGGTTTGAAATTTCGTGGGTGTTTTTACCGGAAGCCCTAATAACAAATCATATCGC 197

Qy       1200 CTGAGCTTAAGAGTTCACNTGTCCAAGCTGCCAAAATGACAAATGAACGCGCGTGGGCAGTT 1259
              ||| |   || | ||| |  ||| |  ||| |  |||| || | | |  || ||   ||||
Db        196 TTTCAAGTAACCCATTCCTGATTTAAACTACCAGTGATGAAGAGAGGATATTTTGAGGAAGTT 137
```

Qy 1260 TATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGA 1319
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 136   TATCAATACCAGTAGGATATTTTTCGAGTAATCAATCACACGAATATTTAGCTGTTGGGATA 77

Qy 1320 CCAAAAAATTCACACAGCTGACAGTCCAGAGCGCCAAAACCGACGACGGCCATATTTTTT 1377
          |||||||||||||||||||||||||||||||||||||||||||||
Db 76   AGAGATAATCCACACAAGAAAGAGACCTGTTTTTTGCCAAGCCCGATGATAGTAATATTTTT 19

RESULT 10
AAT42063_11/c
Continuation (12 of 19) of AAT42063 from base 1100001 (Haemophilus influenzae complete g
WP Sequence split into 19 fragments   LOCUS AAT42063 Accession Aat42063
WP Fragment Name   Begin   End

WP	Fragment Name	Begin	End
WP	AAT42063_00	1	110000
WP	AAT42063_01	100001	210000
WP	AAT42063_02	200001	310000
WP	AAT42063_03	300001	410000
WP	AAT42063_04	400001	510000
WP	AAT42063_05	500001	610000
WP	AAT42063_06	600001	710000
WP	AAT42063_07	700001	810000
WP	AAT42063_08	800001	910000
WP	AAT42063_09	900001	1010000
WP	AAT42063_10	1000001	1110000
WP	AAT42063_11	1100001	1210000
WP	AAT42063_12	1200001	1310000
WP	AAT42063_13	1300001	1410000
WP	AAT42063_14	1400001	1510000
WP	AAT42063_15	1500001	1610000
WP	AAT42063_16	1600001	1710000
WP	AAT42063_17	1700001	1810000
WP	AAT42063_18	1800001	1830121

Query Match 11.6%; Score 160.4; DB 2; Length 110000;
Best Local Similarity 49.6%; Pred. No. 2e-34;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

[Sequence alignment blocks — Qy (query) / Db (database) nucleotide pairs:]

Qy   1   ...   60
Db 104439 ... 104380
Qy   61   ... 120
Db 104379 ... 104332
Qy 121   ... 180
Db 104331 ... 104278
Qy 181   ... 240
Db 104277 ... 104218
Qy 241   ... 300
Db 104217 ... 104158
Qy 301   ... 360
Db 104157 ... 104107
Qy 361   ... 420
Db 104106 ... 104047
Qy 421   ... 480
Db 104046 ... 103987
Qy 481   ... 540
Db 103986 ... 103927

RESULT 11
AAT42063_12/c
Continuation (13 of 19) of AAT42063 from base 1200001 (Haemophilus influenzae complete g
WP Sequence split into 19 fragments   LOCUS AAT42063 Accession Aat42063
WP Fragment Name   Begin   End

WP	Fragment Name	Begin	End
WP	AAT42063_00	1	110000
WP	AAT42063_01	100001	210000
WP	AAT42063_02	200001	310000
WP	AAT42063_03	300001	410000
WP	AAT42063_04	400001	510000
WP	AAT42063_05	500001	610000
WP	AAT42063_06	600001	710000
WP	AAT42063_07	700001	810000
WP	AAT42063_08	800001	910000
WP	AAT42063_09	900001	1010000

[Right column sequence alignment blocks — Qy / Db nucleotide pairs:]

Qy 541 ... 600
Db 103926 ... 103867
Qy 601 ... 659
Db 103866 ... 103807
Qy 660 ... 719
Db 103806 ... 103775
Qy 720 ... 779
Db 103774 ... 103718
Qy 780 ... 839
Db 103717 ... 103658
Qy 840 ... 899
Db 103657 ... 103598
Qy 900 ... 959
Db 103597 ... 103541
Qy 960 ... 1019
Db 103540 ... 103481
Qy 1020 ... 1079
Db 103480 ... 103429
Qy 1080 ... 1139
Db 103428 ... 103388
Qy 1140 ... 1199
Db 103387 ... 103328
Qy 1200 ... 1259
Db 103327 ... 103268
Qy 1260 ... 1319
Db 103267 ... 103208
Qy 1320 ... 1377
Db 103207 ... 103150

```
WP   AAT42063_10                1000000  1110000
WP   AAT42063_11                1100000  1210000
WP   AAT42063_12                1200000  1310000
WP   AAT42063_13                1300000  1410000
WP   AAT42063_14                1400000  1510000
WP   AAT42063_15                1500000  1610000
WP   AAT42063_16                1600000  1710000
WP   AAT42063_17                1700000  1810000
WP   AAT42063_18                1800001  1830121
```

```
Query Match    11.8%;    Score 160.4; DB 2; Length 110000;
Best Local Similarity  49.6%;  Pred. No. 2e-34;
Matches  683; Conservative  0; Mismatches  606; Indels  89; Gaps  7;
```

RESULT 12
AEB39175_30
Continuation (31 of 35) of AEB39175 from base 3000001 (L. pneumophila DNA SEQ ID NO 350°
WP Sequence split into 35 fragments   LOCUS AEB39175 Accession AeB39175
WP Fragment Name   Begin   End

```
WP   AEB39175_00       1         110000
WP   AEB39175_01       100001    210000
WP   AEB39175_02       200001    310000
WP   AEB39175_03       300001    410000
WP   AEB39175_04       400001    510000
WP   AEB39175_05       500001    610000
WP   AEB39175_06       600001    710000
WP   AEB39175_07       700001    810000
WP   AEB39175_08       800001    910000
WP   AEB39175_09       900001    1010000
WP   AEB39175_10       1000001   1110000
WP   AEB39175_11       1100001   1210000
WP   AEB39175_12       1200001   1310000
WP   AEB39175_13       1300001   1410000
WP   AEB39175_14       1400001   1510000
WP   AEB39175_15       1500001   1610000
WP   AEB39175_16       1600001   1710000
WP   AEB39175_17       1700001   1810000
WP   AEB39175_18       1800001   1910000
WP   AEB39175_19       1900001   2010000
WP   AEB39175_20       2000001   2110000
WP   AEB39175_21       2100001   2210000
WP   AEB39175_22       2200001   2310000
WP   AEB39175_23       2300001   2410000
WP   AEB39175_24       2400001   2510000
WP   AEB39175_25       2500001   2610000
WP   AEB39175_26       2600001   2710000
WP   AEB39175_27       2700001   2810000
WP   AEB39175_28       2800001   2910000
WP   AEB39175_29       2900001   3010000
WP   AEB39175_30       3000001   3110000
WP   AEB39175_31       3100001   3210000
WP   AEB39175_32       3200001   3310000
WP   AEB39175_33       3300001   3410000
```

```
WP   AEB39175_34          3400001   3503610

Query Match                    11.4%;   Score  157.6;  DB 14;   Length 110000;
Best Local Similarity          51.9%;   Pred. No. 1.3e-33;
Matches  467; Conservative   0; Mismatches  409;   Indels  24;  Gaps    4;

Qy    170  GCATCACCTTTGAGTAAATCTTGCTCAAATATCGGGGCATCTTGACCGATTAAATACACA  229
           ||||| |||  ||| ||| |   |||||| ||   |||||  ||||||||||||||||||
Db  39768  GGAACTACTTTGGCTAAAGCCGATTCAAATCTTGTCAGCATCTTCACCAATCAAAACAATG  39827

Qy    230  CTATCCACCCACTGCATAACATATCCTGACAGCTCACTAAAATCTTGACCTTTACCAAGC  289
           ||||||||| |||||| | |||||||| |||||||  ||||| |||||||||||||||||
Db  39828  GAACGAACAAAATTCTGAAAACTGGCTAATTCTTGAAAAATCCGACCCCCTTTTCCCTGT  39887

Qy    290  CCACCCAAAAATCAATGCCAGTGATGATATGCCATAAAACCGTGCCAAGCCCATCAATTGCT  349
           |||||| |||||||||||||| ||| |||||||||| |||||||||||||||||||||||
Db  39888  CCACCCGCAATAAGAACAATTTTCCCTTGCATAGACC---TCCTATTCCATTAATAGCA  39944

Qy    350  GCAATGGTTGAGCCGATGGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATA  409
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  39945  GAGATTGTAGCACCAATATTTGTTCCTTTCGAATCATTAAATCCAACCTACTCCATCCACT  40004

Qy    410  TCATCAATATATTCACAGCCATGCGGTAGCCCTTTGAAAGTCTTTAAAGGGTATCAAGCATG  469
           |||||| ||| |||||||||||||||| ||||| ||||| ||||  ||||||||||||||
Db  40005  TCTCTTACCCACTGGCATCGGATGTGGCCAGTCCAGGGAAGGTTTTTAAAACATTAAGAAATG  40064

Qy    470  GCATCCCATGGGAAGAGATTGGCAGCTGTGCTCAAGGCAAGGTGCAGATAAGGCCATTGAGTAGG  529
           |||||||||| ||||| ||| | |||||||| |||||| ||| | |||||||||||||||||
Db  40065  TGTTGCATAGAGAATTCCCGCTGCTTCGGCTTAAAGCAACAGCAAGCCGCATTCATCCAA  40124

Qy    530  TTATGCTTGCCTTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATAT  589
           |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db   530  TTATGCACCCTTTTATTAAAATCGAATCCACAGGCAAAAGACGCTCCATTCCTTCCGCC  40184

Qy    590  AATGCCATACCATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGGTGGCTAAGTAAG  649
           ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  40185  AAATAAGTAGTATTCTTCTTGTTCAATCAGCCCCCAAATTCCCATT--------GAAG  40233

Qy    650  CTATTTGTTGTGTGATTAATGTGTGGGCAACGCTTGTCAGTGTGTGTCAGCATTGC  709
           ||||||||||| |||| ||| | |||| || ||||||||| |||| |||||||||||
Db  40234  GTGGGTCTTTGCCAAAGAGATGCATTTTATATCTGTTTGCACAGGATTGGTGACGAACCG  40293

Qy    710  TTGGCTAAATTGGCATCGTCTTGGTAGATGAATGGGCGGTTGGCAATTTTCAAAAATACGC  769
           ||| |  ||  |||||||||||||||||||||||| ||| |||||||||||||||||||||
Db  40294  T-------ATAGAGGTCTTCTCGATTAAATAAAACAGCCTTCGGCGGTTGGCAAATTCCTT  40346

Qy    770  AATTTTTGTGCCAAATAAATCCACCATGTCCGATCAAGATGATCAGCGCGAGAGA  829
           |||  || | |||||||| ||| ||||| |||||| ||||| |||||||||||||
Db  40347  TGTTTTGCTTGCGGTATAGGCTTCCATAGTCGTGATCCCTGTCAAAATGATCAGGGCTTACA  40406

Qy    830  TTTAAAAATTGTTCGCCCTTGAGCGTGCCAAATTTGAGATAATGCTCAAGCTGAAAAACTTGAC  889
           |||||| ||||||||||||||| |||||||||||||||||||||| |||||||||||||||
Db  40407  TTAAGAATCGTCGCAACCACTGGGCAAGTAAGTTAAATCCAATTGAAAAAACTGGGAT  40466

Qy    890  AGCTCAAGCACCACCAAAATC---CATATTTGGATCGTTAAATATCAAGTGACGTGCAGGCGTG  946
           |||||| |||||||||| |||   |||||| |||  ||||||||| ||||||| ||||||||
Db  40467  AACTCCATACCACAAATCGTAGTGATTCCATCATCATCTAAGACAGGAGTG  40526

Qy    947  CCAATATTACGCGCACCAACAAATCATGCCTGCATGTTTTGCCATCTCGCCCTACTAAT  1006
           |||||| ||||| |||||||| |||||||||||||||||||||||||||||||||||||
Db   947  CCGATATTTCCGGCTACAGCTACCGCGAAACCCGGCGCCCTTAGCCATTTCTCCCACCAAA  40586

Qy   1007  GTGGTAACAGTGCTTTTGGCATTTGAGCCTTGGTGGCGGATGGGTGTTGTTTTGTTGCT  1066
           ||||| |  |||||||||||||||| |||||||||||| ||||||||| |||||||||||
Db  40587  GTCCTTACAGTGGATTTGCCATTAGTTCCCGTTATCGCAATGACAGGGGCGGCTAAATCTCT  40646
```

```
WP   AEB42401_01          100001    210000
WP   AEB42401_02          200001    310000
WP   AEB42401_03          300001    410000
WP   AEB42401_04          400001    510000
WP   AEB42401_05          500001    610000
WP   AEB42401_06          600001    710000
WP   AEB42401_07          700001    810000
WP   AEB42401_08          800001    910000
WP   AEB42401_09          900001   1010000
WP   AEB42401_10         1000001   1110000
WP   AEB42401_11         1100001   1210000
WP   AEB42401_12         1200001   1310000
WP   AEB42401_13         1300001   1410000
WP   AEB42401_14         1400001   1510000
WP   AEB42401_15         1500001   1610000
WP   AEB42401_16         1600001   1710000
WP   AEB42401_17         1700001   1810000
WP   AEB42401_18         1800001   1910000
WP   AEB42401_19         1900001   2010000
WP   AEB42401_20         2000001   2110000
WP   AEB42401_21         2100001   2210000
WP   AEB42401_22         2200001   2310000
WP   AEB42401_23         2300001   2410000
WP   AEB42401_24         2400001   2510000
WP   AEB42401_25         2500001   2610000
WP   AEB42401_26         2600001   2710000
WP   AEB42401_27         2700001   2810000
WP   AEB42401_28         2800001   2910000
WP   AEB42401_29         2900001   3010000
WP   AEB42401_30         3000001   3110000
WP   AEB42401_31         3100001   3210000
WP   AEB42401_32         3200001   3310000
WP   AEB42401_33         3300001   3345687
```

```
Query Match                    11.4%;   Score  157.6;  DB 14;   Length 110000;
Best Local Similarity          51.9%;   Pred. No. 1.3e-33;
Matches  467; Conservative   0; Mismatches  409;   Indels  24;  Gaps    4;

Qy    170  GCATCACCTTTGAGTAAATCTTGCTCAAATATCGGGGCATCTTGACCGATTAAATACACA  229
           ||||| |||  ||| ||| |   |||||| ||   |||||  ||||||||||||||||||
Db  96841  GGAACTACTTTGGCTAAAGCCGATTCAAATCTTGTCAGCATCTTCACCAATCAAAACAATG  96900

Qy    230  CTATCCACCCACTGCATAACATATCCTGACAGCTCACTAAAATCTTGACCTTTACCAAGC  289
           ||||||||| |||||| | |||||||| |||||||  ||||| |||||||||||||||||
Db  96901  GAACGAACAAAATTCTGAAAACTGGCTAATTCTTGAAAAATCCGACCCCCTTTTCCCTGT  96960

Qy    290  CCACCCAAAAATCAATGCCAGTGATGATATGCCATAAAACCGTGCCAAGCCCATCAATTGCT  349
           |||||| |||||||||||||| ||| |||||||||| |||||||||||||||||||||||
Db  96961  CCACCCGCAATAAGAACAATTTTCCCTTGCATAGACC---TCCTATTCCATTAATAGCA  97017

Qy    410  TCATCAATATATTCACAGCCATGCGGTAGCCCTTTGAAAGTCTTTAAAGGGTATCAAGCATG  469
           |||||| ||| |||||||||||||||| ||||| ||||| ||||  ||||||||||||||
Db  97078  TCTCTTACCCACTGGCATCGGATGTGGCCAGTCCAGGGAAGGTTTTTAAAACATTAAGAAATG  97137

Qy    470  GCATCCCATGGGAAGAGATTGGCAGCTGTGCTCAAGGCAAGGTGCAGATAAGGCCATTGAGTAGG  529
           |||||||||| ||||| ||| | |||||||| |||||| ||| | |||||||||||||||||
Db  97138  TGTTGCATAGAGAATTCCCGCTGCTTCGGCTTAAAGCAACAGCAAGCCGCATTCATCCAA  97197

Qy    530  TTATGCTTGCCTTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATAT  589
           |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db  97198  TTATGCACCCTTTTATTAAAATCGAATCCACAGGCAAAAGACGCTCCATTCCTTCCGCC  97257

Qy    590  AATGCCATACCATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGGTGGCTAAGTAAG  649
           ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db  97258  AAATAAGTAGTATTCTTCTTGTTCAATCAGCCCCCAAATTCCCATT--------GAAG  97306

Qy    650  CTATTTGTTGTGTGATTAATGTGTGGGCAACGCTTGTCAGTGTGTGTCAGCATTGC  709
           ||||||||||| |||| ||| | |||| || ||||||||| |||| |||||||||||
Db  97307  GCGGGTCTTTGCCAAAGAGATGCATTTTATATCTGTTTGCACAGGATTGGTGACGA----  97362
```

RESULT 14
ADT05533

```
ID   ADT05533 standard; DNA; 14324 BP.
XX
AC   ADT05533;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Haemophilus influenzae (NTHi) contig DNA sequence - SEQ ID 569.
XX
KW   middle ear bacterial infection; nasopharynx bacterial infection; ds;
KW   contig.
XX
OS   Haemophilus influenzae.
XX
XX
XX
PN   WO2004078949-A2.
XX
PD   16-SEP-2004.
XX
PF   05-MAR-2004; 2004WO-US007001.
XX
PR   06-MAR-2003; 2003US-0453134P.
XX
PA   (CHIL-) CHILDRENS HOSPITAL INC.
XX
PI   Bakaletz LO, Munson RS, Dyer DW;
XX
DR   WPI; 2004-664222/64.
XX
PT   New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT   useful for treating or preventing NTHi bacterial infections of the middle
PT   ear and/or nasopharynx.
XX
PS   Example 1; SEQ ID NO 569; 88pp; English.
XX
CC   The invention comprises nucleotide sequences (genes) from the genome of a
CC   nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC   sequences of the invention are useful for treating or preventing NTHi
CC   bacterial infections of the middle ear and/or nasopharynx. The present
CC   nucleic acid represents an NTHi contig sequence of the invention.
XX
SQ   Sequence 14324 BP; 3171 A; 2271 C; 4310 T; 0 U; 0 Other;
```

Query Match   11.2%;   Score 154.4;   DB 13;   Length 14324;
Best Local Similarity   49.0%;   Pred. No. 4.2e-33;
Matches 591;   Conservative 0;   Mismatches 546;   Indels 69;   Gaps 4;

```
Qy    710 TTGGCTAAATTGGCATCGGTCCTTGGTTAGATGAGATGGGCGGGTTTGGGCAATTTTCAAAAATTAACGC 769
          |||  | ||||  | | | |||||| ||| |||||  |  ||  ||| ||||||    |||||  ||||
Db  97363 ---ACAGTGTAGAGCGTCTTCACGATTAAATAAAACAGCCTTGGGCACCAATTCCTT 97419

Qy    770 TGTTTTGCTTGCGTATAGGCTTCCATAGNTGTAGTCGATGATCAGATCAGCGGAGAGA 829
Db  97420 ...

Qy    830 TTTAAAATTGTTTGCCCCTTGAGCTGTCCAAATTTGAGATATGTCTAAAGCTGAAAAACTTGAC 889
Db  97480 TTAAGAATCGGTCGCGACCAAACCACTGGGCGACAAAAGAATTAAGTTAAATCCAGTTGAAAAACTGGAT 97539

Qy    890 AGCTCAAGCACCACCACCAAAAATC---CATATTTTGGATCGTTAATTAATTCAAGTGCGCAGGCGTG 946
Db  97540 AACTCCAATACCCAACAAAATCGTAGTGTAGTTCATCATCCAAACATATCTAAGACATTAGGTTG 97599

Qy    947 CCAATATTACCGCCGACACCAACAATCATGCCTGCAGTTTTGCCCATTCTGCCTACTAAT 1006
Db  97600 CCGATAATTTCCGGCTACAGCTTACACGAAACCCGCCGGCCTTAGCCCATTTTCCCCACCAAA 97659

Qy   1007 GTCGTAACAGTGCTTTTGGGCATTTGAGTGGGCGATGATTGGGTTGTTGTTTTGTTGCT 1066
Db  97660 GTCGTTACACGTGGGATTTACCATTAGTTCCCGTTATAGCAAATGACAGGGGGCTAAATCTCT 97719
```

(Second alignment column)

```
Qy    172 ATCACCTTTGAGTAAAATCTTGCTCAATTATCGGGGGCATCTTGAACCGATTAAAATACACACT 231
Db   5280 ATAGCTTTGTGATGAAAAATTTTTGCAAGCTGCACACCATCTCGACCAAAAACAATAACAAAT 5339

Qy    232 ATCCACCCACTGCATAACACATATGCTGACAGCTCACTCAAAAAATCTTGACCTTTGACCAAGCCC 291
Db   5340 AATGTGTGGTTGATTAATTCAGCTCATTCTCGAAAAATCAGCCCCTTTTCCGTCC 5399

Qy    292 ACCCAAAATCAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGTGC 351
Db   5400 GCCTAGCAACAAATGCAAT---TTACCCTCAATATATAAAGCCCAGCTTAATGCAGC 5450

Qy    352 AATGGTTGACCGGATGTTGTTGTTACCTTTTGAATCATTAAAAATATGCCTTGCTATCAATATC 411
Db   5451 AACTGTACTCCCACATTGTTTGCTTTAGAGTCATTAATCAACGAATAACCATTACGTTG 5510

Qy    412 ATCAATATATATTCACGCGATGCGGGTAGCCCCTTTGAAAGTCTTAAGGGTATCAAGGCATGGC 471
Db   5511 ATGCACTAATTGAAAAACGATGATCTAACCCTTTGAAAATGACGAAGTGCGGTACGAAGTTG 5570

Qy    472 ATCCATGGGAAGAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTT 531
Db   5571 ATCTAAAATTAATAACCGAACAGCTTGTTGCCCAATGCTGTTGCTCCAAAAAGTTCATATAAATT 5630

Qy    532 ATGCTTGCCTTTGATTTTTAATTGATGGATTGGTGAAAAGTTGTTGTTTTTTTGATGATATAA 591
Db   5631 ATGGCGACCGACTAAGTTACTTCCTTCACAAGGCAAAAATTACTTCATCTTTTACCATTAA 5690

Qy    592 TGCCATACCAATCACCATCATCAAAATAAAATAAAAATCTGCCGTTGGTGGCTAAGTAAGCT 651
Db   5691 GTAATGCTTGCCATTTTCAGTTTTTAGCCAATAATCCGCGACTATATTTCCGCAAAAGAAAC 5750

Qy    652 ATTTGTTGTGATAAGTGTGTGATTTGGGGCAAACGCTTGTCAGTGTGGTCAAGCATTGCTT 711
Db   5751 GGTTGTCTCGCTTGATTTCCACCTTCACCAAAAGTCAGC--- 5790

Qy    712 GGCTAAATTGGCATCGTCTTCTTGGTAGATGAGGTTGGCAATTTCAAAAATACGGCAA 771
Db   5751 ...

Qy    772 TTTTGTTGCCAAAATAATCACCATGCGTGCCGGATCAAGATGATCAGGGAGAGATT 831
Db   5841 TTTTGCTTGGCGGATAATCTTCTAAAATCCATTATAGGATCCATATGGATTCAGTCAGTT 5900

Qy    832 TAAAAATTGTTGTTGCCCCTTGAGCTGCCAAATTTGAGATATGCTAAGCTGAAAAACTTGACAG 891
Db   5901 CAAAAACAGTGCGACCAGCCTTTTAAGCTATATAGTTGTCTCAACGCTAGAAAG 5960

Qy    892 CTCAAGCACCACCAAAATCCATATTTTGGGATCGTTAATTAATTCAAGTGCGTCAGGCGTGCCAAT 951
Db   5961 CTCTAGTACAATAAAGTTCACAATCTTCCATTC---AACAATGACAAAGCGGGAATCCCAAT 6017

Qy    952 ATTACCGCCGACACCAACAATCATCCCTGCAGTTTTGCCATCTCGCCTACTAAAGTCGT 1011
Db   6018 AATTTCCCGCCCATACCAACTTTCACGCCAGAGCTTTTCGGCCATTCATAAAACTAAAGTAGT 6077

Qy   1012 AACAGTGCTTTTGGGCATTTGAGCCTGTGAGCCATGGTGTCTGTTTTGTTGCTTGAGT 1071
Db   6078 AACAGTACATTTTTCCAATCTGTAATCCCACAATTGTTCGCTGCCG--- 6134

Qy   1072 TAGAATCGTCGTTTGACTTTGAGTGTCATCATGAACAAATTGAACATCGCTGACTCAGGAAATGCC 1131
Db   6135 ---------GCAGAATAAATTCAAATATCGCCGGATTACTTCCACTCC 6170

Qy   1132 TTGAGCTTTGGGCGGGGCACAATACTCGGCGTTCTTGGGGTTAAATGCGCGGGCTGATGAT 1191
Db   6171 CGCTTTAAGTACGGTTTGAATTTCTCGGTGTTTTTTACCGGCAAGCCCTGGGCTAATAACAAT 6230

Qy   1192 CTGATCCGGCTGAGCTTTAGAGTTCACTCGTTCCAAGCTGCCAAAATGACAATGAAGGCCTGT 1251
Db   6231 CATATCCGCTTTCAAGTAACCCATTCCTGATTTAAAACTACCAGTATGAAGAGGGATATTTTG 6290

Qy   1252 GGGCAGTTTATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTTGACC 1311
```

```
RESULT 15
ID   ADT05649/c  standard; DNA; 349980 BP.
XX
AC   ADT05649;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Haemophilus influenzae (NTHi) DNA sequence - SEQ ID 685.
XX
KW   middle ear bacterial infection; nasopharynx bacterial infection; ds.
OS   Haemophilus influenzae.
XX
PN   WO2004078949-A2.
XX
PD   16-SEP-2004.
XX
PF   05-MAR-2004; 2004WO-US007001.
XX
PR   06-MAR-2003; 2003US-0453134P.
XX
PA   (CHIL-) CHILDRENS HOSPITAL INC.
XX
PI   Bakaletz LO, Munson RS, Dyer DW;
XX
DR   WPI; 2004-662422/64.
XX
PT   New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT   useful for treating or preventing NTHi bacterial infections of the middle
PT   ear and/or nasopharynx.
PS   Claim 1; SEQ ID NO 685; 88pp; English.
XX
CC   The invention comprises nucleotide sequences (genes) from the genome of a
CC   nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC   sequences of the invention are useful for treating or preventing NTHi
CC   bacterial infections of the middle ear and/or nasopharynx. The present
CC   nucleic acid represents an NTHi DNA sequence of the invention.
XX
SQ   Sequence 349980 BP; 105127 A; 63538 C; 70035 G; 111276 T; 0 U; 4 Other;

Query Match             11.2%;  Score 154.4; DB 13;  Length 349980;
Best Local Similarity   49.0%;  Pred. No. 1.9e-32;
Matches  591; Conservative    0; Mismatches  546; Indels   69; Gaps   4;
```

```
Qy      6291 AGGAAGTTTATCAATACCAGTAGGCTTTTTCGGGTATCAATCACGTATTAGCCTG 6350
             ||||||||
Db    107626 ...

Qy      1312 GTGATTGACCAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAAACCGACGGGCATA 1371
Db    107567 ...

Qy      6351 TTGGGATAAGAGATAATCACACAAGAAAGACCTGTTTTTGCCAAGCCCAATGATCGTAAT 6410
Db    107507 ...

Qy      1372 TTTTTTT 1377
             |||||
Db      6411 ATTTTTT 6416
```

```
Qy       172 ATCACCTTTGAGTAAATCTTGCTCAATTATCGGGGGCATCTTGACGATTAAATACACACT 231
Db    107917 ...

Qy      1072 TAGATCTCGTGCTTTGAGTGTATCAAATGAACAATGAACATGCTGTGATCACAGGAAATGCC 1131
Db    107062 ...

Qy       232 ATCCACCCACTGCATAACATATTGCTCAGCGCTCACTAAAAACTTGACCTTTACCAAGCCC 291
Db    107857 ...

Qy       292 ACCCAAAAATCAATGCCAGTGATGATGAATATGCCAAGCCCATCAATTGCTGC 351
Db    107797 GCCTAGCAACAAAATGCAAT...

Qy       352 AATGGTTGAGCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAAGATC 411
Db    107746 AACTGTACTCCCCACACATTTTGTTGCTTTAGAGTCATTAATAATCCAACGAATACCATTAGCTTG 107687

Qy      1252 GGGCAGTTTATCTGCCAAGGTAGGTAGGAATTGGGGATTACCATGCGATGAGCTAACTTTGTGACC 1311
Db    106906 AGGAAGTTTATCAATACCAGTAGGCTTTTTCGGGTATCAATCACGTATTAGCCTG 106847

Qy      1192 CTGATCGGGCTGAGCTTAAGAGTTCACTGTCTGCCAAAATGAACATGCTGTGATCACAGGAAATGCC 1251
Db    106966 CATATCGGCTTTCAAGTAACCATTCCTGATTAAACTACCAGTATGAAGAGGGATATTTTG 106907

Qy      1132 TTTGAGCTTTGGCGCGACAATACTCGGCGTTCTTGGGGTTTAATGCGGTTGATCGATGAT 1191
Db    107062 ...GCAGAATAATTCAATATGCCGATTACTTCCACTCC 107027

Qy      1012 AACAGTCGCTTTTGGCCATTTGAGCCTGTGAGCGGATGAATGAGGTGGTGGTTGCTTGAGT 1071
Db    107179 ATTTCCGCCCGACACCATTAAACAGACTTTCAGCGACGAGCTTTCCGCCCATTGAGCTAAAGTAGT 107120

Qy       952 ATTACCGCCGACACACCAAAATCATGCCTCCATTCATGTTTTGCCATCTCGCCTACTTGCGT 1011
Db    107236 CTCTTAGTACATAAAGTTCACACAATCTTCATTC---AACAAATGACAAAGCGGGAATCCCAAT 107180

Qy       892 CTCAAGCACCACCCAAAATCCATATATTTGGGATCGTTAATAATTCAAGTGCAGGGGTGCCCAAT 951
Db    107296 CAAAACAGTTGCACGCCGCAGCTTTTTTAAGCTATAAGTTGTTCTCAAGCTGAAAACTAGAAAG 107237

Qy       832 TAAAATTGTTGTTGCCCCTTTGGAGCTGCCGAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAG 891
Db    107356 TTTTTTGCTTGGCGAGTAATCTTCTAAAATCCATATAGGGGAT... 107297

Qy       772 TTTTTGTGCAAATAATCCACCATGTGCGGATCAAGATAATCAGCCGGGAGAGATT 831
Db    107406 --------------TTATCCTCATTATTCAAAAACTCCGACCTTAGCTAAAATGCGTAA 107357

Qy       712 GGCTAAATTGGCATGCTGTCTTGGTAAGATGATGAGGCCGGTTTGGCAATTTTTCAAAAAATACGCAA 771
Db    107446 GGTTTGGTCTCGGCTTAATTCACCTTCAACCCAAAAAGTCAGC-------- 107407

Qy       652 ATTTGTTGTGATAATGGTGTTGAAGTGGTGCCGATTTTTTAGCCGCAAATAAATCCGCGTTTGGTGGC... 711
Db    107506 GTAATGCTTGCCATTTTCAGTTTTTAGCCAATAAATCCGACTATTTTCCGGCAGACATTGCTT 107447

Qy       592 TGCCATACCATCACCATCACAAATAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCT 651
Db    107566 ATGGCGGACCCGACTAATGTTACTTCTTCACAAGGCAAAAATTACCTTCATCTTTTACCATTAA 107507

Qy       532 ATGCTTGCCTTTGATTTTTTAATTGATGGATTGGTAAAAGTTGTTTTTGTTTTTTGATGATATAA 591
Db    107626 ATCTAAAAATTAATAACCGACAGCTTGTTGCCAAATGCTTGTGCCAAATGATCGTAAT 107567

Qy       472 ATCCATGGGGAGAGATTGGGCAGCTGGCCTAAGGCAGAGTAGCAGTGCGAGAGTATGAGGTAGGTT 531
Db    107686 ATGCACTAATTGAAAACGATGAATCTAAACCCTTTGAAAATGACGAAGTGCGGTTACGGAAATTGA 107627

Qy       412 ATCAATATATTCACAGCGATGCGATGCGTAGCCCTTGAAAGTCTTAAGGGTATCAAGCATGGC 471
Db    107746 ...
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: May 12, 2006, 07:39:28 ; Search time 7220 Seconds
                        10864.815 Million cell updates/sec
                        (without alignments)

Title: US-10-672-787-35_COPY_11357_12736
Perfect score: 1380
Sequence: 1 taaagcgtgaaccatgttaa.........acgacgcatattctttga 1380

Scoring table: IDENTITY NUC
              Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 2842172S653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                Maximum Match 100%
                Listing first 45 summaries

Database :   GenEmbl:*
           1: gb_ba:*
           2: gb_in:*
           3: gb_en:*
           4: gb_env:*
           5: gb_om:*
           6: gb_ov:*
           7: gb_pat:*
           8: gb_ph:*
           9: gb_pr:*
         10: gb_ro:*
         11: gb_sts:*
         12: gb_sy:*
         13: gb_un:*
         14: gb_vi:*
         15: gb_htg:*
         15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	DB ID	Description
1	1380	100.0	96109	6	AR408756	AR408756 Sequence
2	1380	100.0	96109	6	AX067460	AX067460 Sequence
3	1370.4	99.3	1410	6	AR450009	AR450009 Sequence
4	266.8	19.3	110000	1	CP000082_25	Continuation (26 0
5	242	17.5	110000	1	CR543861_02	Continuation (3 of
6	229.6	16.6	110000	1	CP000082_24	Continuation (25 0
7	226.2	16.4	1416	1	AR318732	AR318732 Sequence
8	160.4	11.6	1314	6	AX932204	AX932204 Sequence
9	160.4	11.6	11498	1	U32793	U32793 Haemophilus
10	160.4	11.6	11498	1	AX191762	AX191762 Sequence
11	160.4	11.6	110000	6	BD426631_11	Continuation (12 0
12	160.4	11.6	110000	6	BD426631_12	Continuation (13 0
13	160.4	11.6	110000	6	AR274513_11	Continuation (12 0
14	160.4	11.6	110000	6	AR274513_12	Continuation (13 0
15	160.4	11.6	110000	6	AR632719_11	Continuation (12 0
16	160.4	11.6	110000	6	AR632719_12	Continuation (13 0
17	157.6	11.4	110000	1	CR628336_30	Continuation (31 0
18	157.6	11.4	110000	1	CR628337_28	Continuation (29 0
c 19	154.8	11.2	110000	1	AE016828_01	Continuation (2 of
c 20	154.4	11.2	14324	1	CQ872956	CQ872956 Sequence
c 21	154.4	11.2	110000	1	CP000057_12	Continuation (13 o
c 22	154.4	11.2	349980	6	CQ873072	CQ873072 Sequence
c 23	152.8	11.1	110000	1	AE017354_29	Continuation (30 o
c 24	152.4	11.0	1314	6	AX932206	AX932206 Sequence
c 25	146.2	10.6	110000	1	BA000031_04	Continuation (5 of
c 26	142.2	10.3	12232	1	AE006048	AE006048 Pasteurel
c 27	139.6	10.1	1317	6	AX189058	AX189058 Sequence
c 28	139.6	10.1	28277	1	BC2MIN	BC2MIN
c 29	139.6	10.1	28277	6	AX191720	AX191720 Sequence
c 30	139.6	10.1	2608	1	U00096_00	ECMUROY
c 31	139.6	10.1	111408	1	ECO110K	D10483 Escherichia
c 32	139.6	10.1	300409	1	ECO110K	D10483 Escherichia
c 33	139.6	10.0	110000	1	AE016755	AE016755 Escherich
c 34	137.4	10.0	110000	1	AE017340_04	Continuation (5 of
c 35	134.8	9.8	110000	1	AE005174_00	Continuation (1 of
c 36	134.8	9.8	110000	1	AE005174_01	Continuation (2 of
c 37	134.8	9.8	110000	1	BA000007_00	Continuation (1 of
c 38	134.8	9.8	110000	1	BA000007_01	Continuation (2 of
c 39	133.2	9.7	1629	1	X17609	X17609 Escherichia
c 40	131.6	9.5	110000	1	AE005674_00	Continuation (1 of
c 41	131.6	9.5	290029	1	AE016978	AE016978 Shigella
c 42	130.2	9.4	110000	1	AE016827_16	Continuation (17 o
c 43	125.2	9.1	1323	6	AR377824	AR377824 Sequence
c 44	124.6	9.0	110000	1	CP000058_46	Continuation (47 o
c 45	124	9.0	304558	1	AE017151	AE017151 Haemophil

ALIGNMENTS

RESULT 1
LOCUS   AR408756   96109 bp   DNA   linear   PAT 18-DEC-2003
DEFINITION Sequence 35 from patent US 6632636.
ACCESSION   AR408756
VERSION   AR408756.1   GI:40159151
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM Unknown.
           Unclassified.
REFERENCE   1 (bases 1 to 96109)
  AUTHORS   Lagace,R.E., Patterson,C. and Berg,K.L.
  TITLE   Nucleic acids encoding 3-ketoacyl-ACP reductase from Moraxella
           catarrhalis
  JOURNAL   Patent: US 6632636-A 35 14-OCT-2003;
           Eiitra Pharmaceuticals Inc.; San Diego, CA
FEATURES   Location/Qualifiers
    source   1..96109
                      /organism="unknown"
                      /mol_type="genomic DNA"

ORIGIN

Query Match   100.0%; Score 1380; DB 6; Length 96109;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTCATTATATTTAAATTGATC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11357 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTCATTATATTTAAATTGATC 11416

Qy     61 CATGCTAGCACAAGCAGGTGACGAAAAGCACAGCCCTAGCCGCGTGCTTGACTGGCTGCCAA 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11417 CATGCTAGCACAAGCAGGTGACGAAAAGCACAGCCCTAGCCGCGTGCTTGACTGGCTGCCAA 11476

Qy    121 TACTGCCCGATCAAGTGTACCAAACCTTGATATATTTTAGATGCCAAATGTGCATCACCTTT 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11477 TACTGCCCGATCAAGTGTACCAAACCTTGATATATTTTAGATGCCAAATGTGCATCACCTTT 11536

Qy    181 GAGTAAAATCTTGCTCAATTATCGGGGGCCATCTTGACGGATTAAAATACACACTATCCACCCA 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11537 GAGTAAAATCTTGCTCAATTATCGGGGGCCATCTTGACGGATTAAAATACACACTATCCACCCA 11596
```

RESULT 2

```
LOCUS       AX067460    96109 bp    DNA    linear    PAT 24-JAN-2001
DEFINITION  Sequence 35 from Patent WO0078968.
ACCESSION   AX067460
VERSION     AX067460.1  GI:12545080
KEYWORDS
SOURCE      Moraxella catarrhalis
  ORGANISM  Moraxella catarrhalis
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Moraxellaceae; Moraxella.
REFERENCE   1
  AUTHORS   Legace,R.E., Patterson,C. and Berg,K.L.
  TITLE     Nucleotide sequences of moraxella catarrhalis genome
  JOURNAL   Patent: WO 0078968-A 35 28-DEC-2000;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..96109
                     /organism="Moraxella catarrhalis"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:480"
ORIGIN

Query Match             100.0%;  Score 1380;  DB 6;  Length 96109;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches  1380;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

RESULT 3

```
LOCUS        AR450009/c              1410 bp    DNA         linear    PAT 20-FEB-2004
LOCUS        AR450009
DEFINITION   Sequence 673 from patent US 6673910.
ACCESSION    AR450009
VERSION      AR450009.1  GI:42679066
KEYWORDS     .
SOURCE       Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 1410)
  AUTHORS    Breton,G.L.
  TITLE      Nucleic acid and amino acid sequences relating to M. catarrhalis
             for diagnostics and therapeutics
  JOURNAL    Patent: US 6673910-A 673 06-JAN-2004;
             Genome Therapeutics Corporation; Waltham, MA
FEATURES             Location/Qualifiers
     source          1..1410
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN
```

Query Match   99.3%; Score 1370.4; DB 6; Length 1410;
Best Local Similarity   99.6%; Pred. No. 0; Mismatches 0;
Matches 1374; Conservative   0; Indels 0; Gaps 0;

```
Qy       1 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATATATTTTAAATTGATC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11957 ATCACCATCATCAAATAAATAAAAATCTGACCGTTTGTGTGGCTAAGTAAGCTATTGTTGT 12016

Qy      61 CATGCTAGCACAAGCAGGTGACAAAAAGCACAGCACCCCTAGCGGTGCTTTGACTGCCGCC 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12017 GATAATGGTGTGATTGGGGCAACGGCTTGTCAGTGTGGTCAAGCCATTGCTTGGCTAAATT 12076

Qy     121 TACTGCCCGATCAAGTGTACCAAACCTGATATATTTTAGATGCCAAAATGTGCATCACCTTT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12077 GGCATCGTCTTGGTAGATGATGGGCGGTTTGGCAATTTTCAAAAATACGCAAATTTTGTGC 12136

Qy     181 GAGTAAATCTTGCTCAATTATCCGGGGCATCTTGACGGATTAAATACACACTATCCAATA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12137 CAAATAATCCACCATCGTGCCCGATCAAGATGATCAGCGGAGGAGAGGATTTAAAATTGT 12196

Qy     241 CTGCCATAACATATGCTGACAGCTCACTAAAATCTTCACCTTTACCAAGCCCCACCAAAAT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12197 TGCCCCCTTGAGCTGCCAAATTTGACAGCTCAAGCAC 12256

Qy     301 CAATGCCAGTGATGATGATGATGCCATAAAACCGTGCCAAGCCCCATCAATTGCTGCAGTT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12257 CACCAAATCCATTATTTTGGGATCGTTAATAATTCAAGTGCAGGCGGTGCCAATATTACCGCC 12316

Qy     361 GCCGATGTTTGTACCTTTTGAATCATTAAAAATATGCCTTGCTGCTATCAATATCATCAATATA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12317 GACACCAACAATCATCGCTGCATGTTTTGCCATCTCGCCTACTAAGTCGTAACAGTGCT 12376

Qy     421 TTCACACGCGATCGCGGTAGCCCCTTTGAAGTGTCTTAAGGGTATCATGGCATCCATGGG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12377 TTTGGCATTTGAGCCGTGTGATGGCGATGATTGGTGTTGTTTGTTTGCTTGTGAGATCTCG 12436

Qy     481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCAATGAGTGAGTTATGCTTGCC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12437 TGCTTTGAGTGTATCAAATGACAAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTTT 12496

Qy     541 TTTGATTTTTAATTGATCGGATTCGTAAAAGTTGTTGATGATGATAATGCCATACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12497 GGCGCGGGCAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGGC 12556

Qy     601 ATCACCATCAAATAAAATAAAATCTGCCGTTTGGTGGGCTAAGTAAGCTATTGTTGT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12557 TGAGCTTAAGAGTTCACTGTGCCAAAATGAACGCCCTGTGGGCAGTTT 12616

Qy     661 GATAATGGTGTGATTGGGGCAACGGCTTGTCAGTGTGGTCAAGCCATTGCTTGGCTAAATT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12617 ATCTGCCAAGGTAGGATTGGGATTACCATCAAGATGATGTAACTTGTGACCCTGATTGAC 12676

Qy     721 GGCATCGTCTTGGTAGATGATGGGCGGTTTGGCAATTTTCAAAAATACGCAAATTTTGTGC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12677 CAAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACCGGCCATATTTTTTGA 12736

Qy     781 CAAATAATCCACCATCGTGCCCGATCAAGATGATCAGCGGAGGAGAGGATTTAAAATTGT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CAAATAATCCACCATCGTGCCCGATCAAGATGATCAGCGGAGGAGAGGATTTAAAATTGT 840

Qy     841 TGCCCCCTTGAGCTGCCAAATTTGACAGCTCAAGCAC 900
           ||||||||||||||||||||||||||||||||||||
Db     841 TGCCCCCTTGAGCTGCCAAATTTGACAGCTCAAGCAC 900

Qy     901 CACCAAATCCATTATTTTGGGATCGTTAATAATTCAAGTGCAGGCGGTGCCAATATTACCGCC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CACCAAATCCATTATTTTGGGATCGTTAATAATTCAAGTGCAGGCGGTGCCAATATTACCGCC 960

Qy     961 GACACCAACAATCATCGCTGCATGTTTTGCCATCTCGCCTACTAAGTCGTAACAGTGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GACACCAACAATCATCGCTGCATGTTTTGCCATCTCGCCTACTAAGTCGTAACAGTGCT 1020

Qy    1021 TTTGGCATTTGAGCCGTGTGATGGCGATGATTGGTGTTGTTTGTTTGCTTGTGAGATCTCG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TTTGGCATTTGAGCCGTGTGATGGCGATGATTGGTGTTGTTTGTTTGCTTGTGAGATCTCG 1080

Qy    1081 TGCTTTGAGTGTATCAAATGACAAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTTT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 TGCTTTGAGTGTATCAAATGACAAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTTT 1140

Qy    1141 GGCGCGGGCAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGGC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GGCGCGGGCAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGGC 1200

Qy    1201 TGAGCTTAAGAGTTCACTGTGCCAAAATGAACGCCCTGTGGGCAGTTT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 TGAGCTTAAGAGTTCACTGTGCCAAAATGAACGCCCTGTGGGCAGTTT 1260

Qy    1261 ATCTGCCAAGGTAGGATTGGGATTACCATCAAGATGATGTAACTTGTGACCCTGATTGAC 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 ATCTGCCAAGGTAGGATTGGGATTACCATCAAGATGATGTAACTTGTGACCCTGATTGAC 1320

Qy    1321 CAAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACCGGCCATATTTTTTGA 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CAAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACCGGCCATATTTTTTGA 1380
```

```
RESULT 4
CP000082_25
WPCOMMENT
WPCOMMENT
Sequence split into 27 fragments   LOCUS CP000082 Accession CP000082
Fragment Name     Begin      End
CP000082_00            1      110000
CP000082_01       100001      210000
CP000082_02       200001      310000
CP000082_03       300001      410000
CP000082_04       400001      510000
CP000082_05       500001      610000
CP000082_06       600001      710000
CP000082_07       700001      810000
CP000082_08       800001      910000
CP000082_09       900001     1010000
CP000082_10      1000001     1110000
CP000082_11      1100001     1210000
CP000082_12      1200001     1310000
CP000082_13      1300001     1410000
CP000082_14      1400001     1510000
CP000082_15      1500001     1610000
CP000082_16      1600001     1710000
CP000082_17      1700001     1810000
CP000082_18      1800001     1910000
CP000082_19      1900001     2010000
CP000082_20      2000001     2110000
CP000082_21      2100001     2210000
CP000082_22      2200001     2310000
CP000082_23      2300001     2410000
CP000082_24      2400001     2510000
CP000082_25      2500001     2610000
CP000082_26      2600001     2650701

Continuation (26 of 27) of CP000082 from base 2500001 (CP000082 Psychrobacter arcticum 2

Query Match        19.3%;  Score 266.8;  DB 1;  Length 110000;
Best Local Similarity  53.5%;  Pred. No. 7.1e-54;
Matches  745;  Conservative  0;  Mismatches 572;  Indels 75;  Gaps  6;
```

[Rotated sequence alignment data — pairwise Qy/Db nucleotide sequence blocks with position numbers, e.g.:]

```
Qy   1021  TTTGGCATTTGAGCCCTGTGATGGCCATGATTGGGTGTGTTGTTGCTTGAGTTAGAATCTCG  1080
Db    372  TTTGGGCATTTGAGCCCTGATGATGGCGATGATTGGGTGTGTTGTTGCTTGAGTTAGAATCTCG  313

... (sequence alignment continues) ...

Qy    128  CGA-----------...---------TCAAGTGTACCAACCTGATATATATTTTAGATGCCAAAATGT  169
Db   9997  ATTATAAATCCGTGCCGCTTCTTGTCATCGGATCGGACCAAAACTGTCTTGATA  10056
```

RESULT 5
CR543861_02/c

WPCOMMENT
Sequence split into 36 fragments   LOCUS CR543861   Accession CR543861

Fragment Name	Begin	End
CR543861_00	1	110000
CR543861_01	100001	210000
CR543861_02	200001	310000
CR543861_03	300001	410000
CR543861_04	400001	510000
CR543861_05	500001	610000
CR543861_06	600001	710000
CR543861_07	700001	810000
CR543861_08	800001	910000
CR543861_09	900001	1010000
CR543861_10	1000001	1110000
CR543861_11	1100001	1210000
CR543861_12	1200001	1310000
CR543861_13	1300001	1410000
CR543861_14	1400001	1510000
CR543861_15	1500001	1610000
CR543861_16	1600001	1710000
CR543861_17	1700001	1810000
CR543861_18	1800001	1910000
CR543861_19	1900001	2010000
CR543861_20	2000001	2110000
CR543861_21	2100001	2210000
CR543861_22	2200001	2310000
CR543861_23	2300001	2410000
CR543861_24	2400001	2510000
CR543861_25	2500001	2610000
CR543861_26	2600001	2710000
CR543861_27	2700001	2810000
CR543861_28	2800001	2910000
CR543861_29	2900001	3010000
CR543861_30	3000001	3110000
CR543861_31	3100001	3210000
CR543861_32	3200001	3310000
CR543861_33	3300001	3410000
CR543861_34	3400001	3510000
CR543861_35	3500001	3598621

Continuation [3 of 36] of CR543861 from base 200001 (CR543861 Acinetobacter sp. ADP1 com

Query Match   17.5%;   Score 242;   DB 1;   Length 110000;
Best Local Similarity   52.9%;   Pred. No. 7.2e-48;
Matches 726;   Conservative   0;   Mismatches 580;   Indels 66;   Gaps 7;

Qy   5 GCGTGAACCATGTTAACAAACTTATCACCTTCGCTCATTATATAATTTTTTAAATTGATCCATG 64
            ||| ||||||| || || ||||||||||||||||||| || |||| ||||||||||||||
Db   71434 GCATCCACACAGGCAACAAAATTGAGTGTCCCGATCTCGGATTAACCTTTTTAAACATATCAAA 71375

Qy   65 CTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCCGGTCGCTTTGACTGGCTGCCAATACT 124
            ||||||||||||||||| || ||||||||| ||||||| |||||| ||||| |||| |||||
Db   71374 ATCGCTTGCGTCGGATTTCGGGTTAAGTGTTGTGCCAGTCCAGTTAAAAATATCGATCGAGT 71321

Qy   125 GCCCGATCAAGTGTACCAACCTTGATATATTTTAGATGCCAAAATGTGCATCACCTTTGAGT 184
            ||| | ||| || ||||| || |||| ||||||||||||||| ||| ||| ||| |||||
Db   71320 GCCTGTGTCACCGCCATCTTTTAAGTGATGCATCGCACCAACTGAGTTGTTCTGT 71264

Qy   185 AAATCTTGCTCAATTTATCGGGGCGCATCTTCTTTGACCGATTAAAATACACATTCCACCACTGC 244
            ||| || | ||| | |||||||| |||||||||| ||||| || ||||| |||||| ||||
Db   71263 AACGCCTGCTCAAATCTGCTGAGTTTAATCTGTTAATCGGGGATCTTCACGCAGAAATCGGCATCA 71204

Qy   245 ATAACATATGCTGACAGCTCACTATAAAATCTTCGACCTTTACCAAAGTCGGCACCCCAAAATC--- 301
            |||| || | || |||||||||| |||||||| ||||||||| |||| |||||||||||||||
Db   71203 TCAACAGACTGGCGCAGCAGAGTCAAGAGTAAATCCCTGGCCTTTTACCCTGGCGCCCAAAAACACA 71144

Qy   302 ---AATGCCAGTGATGAATATGCCATAAACGTGCCAAGCCCATCAATTGCTCGCAATTGTT 358
               |||||||| | | || ||||||||| ||||||||||||||||| | | |||||||||
Db   71143 ATGACTTTTGCCTTGCTTGTTTGGGTTCAAATTGCAGCACCAAGACCATCAATCGGCAAGCTTT 71084

Qy   359 GAGCCGATGTTTGTACCTTTTTGAATCATTACATTAAAAATATGCCTTGTCTATATACATCAAATA 418
            ||||| ||||| || |||||| ||||| |||||||||||| |||||||||||||| |||||||||
Db   71083 GCCCCCACATTTGGTCCCTTTCCGAATCATTATAAACGTACATCATTGACTGTTTTTTTPACA 71024

Qy   419 TATTCGCAGCGATTGGCGGTAGCCCTTTGAAAGTCTTAAGGGGTATCAAGCCATGATCGACA 478
            ||||||||||||| || |||||||||||||||||||||||| |||||||||||||| ||||
Db   71023 TATTCCGCAGCGATTGCTCAAAGCGCCTTTAAACTGCTTCAAGGTTTCAGCCATCAAGCCATCATGTCTACA 70964

Qy   479 GGAAGATTGGCAGCTGGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTG 538
            |||||||||| |||||||||| ||||||||| ||||||||| |||||||| ||||||||
Db   70963 GGTAAACCTATCGCTTCACCCAAATGCCAGAACAAAGCCAGTGCAATATATTATGTGTG 70904

Qy   539 CCTTTGATTTTAATTGATGGATTGGTAAAAGTTGTTTTTTTGGATGATATAAATGCCATA 598
            ||||||| | ||||| |||||||||||||||||||||||||||||||||||||| |||
Db   70903 CCCTGAATATACAGCTCAGAACTCTTAATCAGGGCGGATACGGCCAGCCACATG 70844

Qy   599 CCATCACCATCATCAAAATAAAAAATCGCGTTTGGGTGGCTAAGTAAGTTATTTGTT 658
            ||||||||| |||||| ||| ||||||||| | |||| || |||||||||| ||||
Db   70843 GTTCCATCATCTTCACGTAGTAAAAACGCCATATGCTT-----AAATGTCGG 70801

Qy   659 GTGATAATGGTCTGTGATTTGGGGCCAACGGCTTGTCTCCAGTGTGGTCAGTTGTCAAAC 718
            |||||||||||| | |||||| |||||| || ||||||||| || ||| ||||| ||
Db   70800 GTGCATTGAGTCCAAAAACTTGCATCGGTTGTCTGTACTATCTGGCGACCAGAGGGTGTCAAAC 70741

Qy   719 TTGGCATCGTCTCTTGGTTAGATGAATGGCGGGTTTGGCAAGTTTGGCAATTTTGT 778
            |||||| |||||| | | ||||||||||| | || |||||| |||| |||||||
Db   70740 T----GTCATCACGGTTGTAACCACCTTTTTTACACCTGAAAAAAATACGAAATTTACC 70685

Qy   779 GCCAAATAATCCACCATGCCATCGTGCCGATCAAGATGAATGAGGCGGAGAGATTAAAATT 838
            || |||||| |||| || || || |||||||||| ||||| |||| || ||||| ||||||
Db   70684 TGATGATAGCCCAGCATGTCTCCATGACGGATCTAGATGAATGCCTCACTCATTATTTAATATG 70625

Qy   839 GTTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTCAAGCTCAGC 898
            ||||| |||||||||||||||||||| |||| |||||||||||||||||||
Db   70624 ACTGCCACCTCAGCATTCAGGTTGCGATTCAGGTCGTTTCCAGTGAAAAACTTGACAGCTCTAGT 70565

Qy   899 ACCACCAAAATCCATATATTTTGGATCGTGTTAAATTCAAGTCAGAGGGGTGCCAATATTACCG 958
            |||||||| ||||| ||| || ||||||||||||||||||| || |||| |||| || ||||
Db   70564 ACCAGCAACTC---AGCCTCGTCTTTTTAACAAATCTAATCCTGGACGGGCCCTAAAATTACT 70508

Qy   959 CCGACACCAACAAATCATGCCATGTTTGCCATCTTTCGCCTACTGTTGTAAACAGTG 1018
            || || ||||| || ||||| ||||||||||||| | ||||||||| ||| |||||
Db   70507 CCTACTGCCACGCGCGTTTACCTGCCATCTTTGCCATCTTTCGACCTTACTCTACTGAGTTGTTTACGGTA 70448

Qy   1019 CTTTTGCATTTGAGTGTATCAATGAACATCGCTGATGGCCATGATTGGTCGTGTTTGTTTGCTTTGAGTTAGATCT 1078
            ||||||||||||| | | ||||||||||||||| ||||| ||||| |||||| | ||| |||||||||| |
Db   70447 CTTTTTGGCATTTGAACCCGTAATCGGAACAAATCGGAACCTGAGTTGTCG-----CTTTTTTCT--- 70394

Qy   1079 CGTGCTTGAGTGTATCAATGAACATCGCTGATCGACAGGAATGCCTTGTGCA 1138
            ||| ||||| ||||||| |||| | || | |||| || || || | |||
Db   70393 ------------------AAGACTGATATCTCCATGACTGGAATGCCTTGTGCA 70355

Qy   1139 TTGGCGGCGACAATACTCGGCGGTTCTTGGGGTTAATGCCGGTTAATGGACAGATGATCGAG 1198
            |||| || | ||||||||| |||||||||| |||| ||||||||||| ||| |||||||
Db   70354 ATCGCTTGCGTCGGATTTCGGGTTAAGTGTTGTGCCAGTCCAGTTAAAAATATCTTCG 70295

Qy   1199 GCTGAGCTTAAGAGTTCACTCGTCCAAAGCTGCCAAAATGACAAATGAACGCCGTGGGCAGT 1258
            ||||||||||| || |||| | |||| || || ||| || ||| |||| | |||||||
Db   70294 GCTTGCGAGGAGTAGTTCAGTTATCTAACTGACCAAAAACTGGCTTTTGCACATCTTGAGGAATC 70235

Qy   1259 TTATCTGCCAAGGTAGGATTGGGATTACCATCGATGAGTGCTAACTTGTGACCGTGAATTG 1318
            ||||||||||||||| |||| | |||||||||||||||||| ||| |||||| |||| |||
Db   70234 ...

Qy   1361 ACGACGGCATAT 1372
            ||||||||||||
Db   10177 ACCACGACTTGT 10188

Qy   1301 ACTTCTTGTGACCGTGAGATTGACCAAAAAATTCACAGCTGATAGACGACTCGCCCAAACCG 1360
            |||||| | |||||||||||||||||||||||| || ||||||| ||| || |||||||||
Db   10117 ACTTCGTAGCCCTGCTCAGCCAAATAATGTCCGACCGATAATCCCGACTGCCCAAAACCG 10176

Qy   1241 TGAACGCCTGTGGGCAGTTAICTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTA 1300
            |||||||||| ||||||| | ||||||||||||||||||||||||||||||||||||
Db   10057 TCAATAGTAGCAGGCAGTTTATCTGCTTAATCGCGATTCTCCTGATTGTCAGTGACGGCA 10116

```
Db  70234  TGATCAATGCCCAGTGGGAGTCTCTGAGAATCTGTTGACAGCAACGGCGGTAGCCTTGGTTCA 70175
            ||||                |||||||           |||||||  |||||||
Qy   1319  ACCAAAAAATTCACAGCTGACAGTCCAGAGCGTGCCCAAACCGACGACGGCAT 1370

Db  70174  TGTAAAAAATTAACGGGTCGAAACCCCTGAGATTCCAAGACCCGCAACGGACTT 70123

RESULT 6
CP000082_24
WPCOMMENT
Sequence split into 27 fragments    LOCUS  CP000082  Accession CP000082

Fragment Name       Begin      End
CP000082_00         110001     110000
CP000082_01         100001     210000
CP000082_02         200001     310000
CP000082_03         300001     410000
CP000082_04         400001     510000
CP000082_05         500001     610000
CP000082_06         600001     710000
CP000082_07         700001     810000
CP000082_08         800001     910000
CP000082_09         900001     1010000
CP000082_10         1000001    1110000
CP000082_11         1100001    1210000
CP000082_12         1200001    1310000
CP000082_13         1300001    1410000
CP000082_14         1400001    1510000
CP000082_15         1500001    1610000
CP000082_16         1600001    1710000
CP000082_17         1700001    1810000
CP000082_18         1800001    1910000
CP000082_19         1900001    2010000
CP000082_20         2000001    2110000
CP000082_21         2100001    2210000
CP000082_22         2200001    2310000
CP000082_23         2300001    2410000
CP000082_24         2400001    2510000
CP000082_25         2500001    2610000
CP000082_26         2600001    2650701
CP000082_26         2650001    2650701

Continuation [25 of 27] of CP000082 from base 2400001 (CP000082 Psychrobacter arcticum 2

Query Match         16.6%;    Score 229.6;  DB 1;   Length 110000;
Best Local Similarity  54.3%;   Pred. No. 7.2e-45;
Matches 581;   Conservative  0;  Mismatches 444;  Indels 45;  Gaps 4;

Qy     8  TGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATCATGCTA 67
          |||||||||      ||||| ||||||||||||||||||      ||||||||||||
Db 108845 TTAACGAGCTGGGCTAAAGCGGCTCGCCCCGCGCAAACCCATTAAACTGATCGAAGCTT 108904

Qy    68  GCACAAGCAGCGAGGTGGACAAAAGCACAGCCCCTTAGGCGGTGCTTGACTGGCTGCCCATTAAACTACTTCGGTTTACTTGC 127
          |||||||||       |||||||||        |||||||||||||      ||||
Db 108905 GCACAAGCAGGTGACAATAATACCGCCCTGTACTTGCGACAAAACTACTTCGGTTTACTTGC 108964

Qy   128  CGA------------------------------TCAAGTGTACCAACCTGATATATTTAGATGCCAAAATGT 169
                                          |||||||||||||||||||||||||
Db 108965 TCAATCATCGAAAAAGCATTTTCTAAGGTTCTGACATTGAATGCAGGCTAACCTCATCGGCTG 109024

Qy   170  GCATCACCTTTGAGTAAATCTTGCTCAATTATCGGGCANCTTGACCGATTAAATAACACA 229
          |||||||       ||||||||||      |||||||||
Db 109025 AGACCCGCTGCGCCTAAATGCTGTTCAATCTGACTTGTCCTCGCCAATAAAGAGCACT 109084

Qy   230  CTATCCACCCACCTGCATAACATATGCTAACAAACTTGACCTTTACCAGC 289
          |||||          |||||||||||||||
Db 109085 TGACTCACATATTGATTGATAAAAGGCGAGAGTTCACCGAACTGCTTACCTTGA 109144

Qy   290  CCACCCAAAATCAATGCCA------------GTGATGATATGCCATAAACGGTGCCAAGCCCATCA 343
          |||||||||||||||||||            ||||||||||||||||
Db 109145 CCGCCCAATATTAATAAACTTACCGTCTTTGGGGTGCATAGACCGCGCCCAAACCTTCA 109204

Qy   344  ATTGCTGCAAATGGTTGAGCGCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTA 403
          |||||||    ||||||||||||||||||||||||||||      |||||||||
Db 109205 ACAGCCGCGCCATGGTTGAGCCAAATATTGGTGCCTTTAGAATCATTAAAGTAATCAAATGCCA 109264
```

```
RESULT 7
AR318732/c
LOCUS       AR318732                 1416 bp    DNA     linear   PAT 17-AUG-2003
DEFINITION  Sequence 1282 from patent US 6562958.
ACCESSION   AR318732
VERSION     AR318732.1  GI:3369835
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1416)
  AUTHORS   Breton,G. and Bush,D.
  TITLE     Nucleic acid and amino acid sequences relating to Acinetobacter
            baumannii for diagnostics and therapeutics
  JOURNAL   Patent: US 6562958-A 1282 13-MAY-2003;
            Genome Therapeutics Corporation; Waltham, MA
FEATURES             Location/Qualifiers
     source          1..1416
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match         16.4%;    Score 226.2;  DB 6;   Length 1416;
Best Local Similarity  52.1%;   Pred. No. 8.6e-44;
Matches 718;   Conservative  0;  Mismatches 593;  Indels 66;  Gaps 7;

Qy   404  TCAATATCATCAATATAATTCACAGGATGCGGTGTAGCCCTTTGAAAGTCTTAAAGGGGTATCA 463
Db 109265 GCAATTGTGGCGACATATTGGCGACGGGATGCTCAAGACCTCACAAACTGCTA 109324

Qy   464  AGCATGGCATCCATCATGGGAAGAGTTGGCAGCTGTGCCTAAGCGAAGTGCAGATAAGGCATTG 523
Db 109325 AGCATACTTCCAAGCGGCAAAACACAGCAAGCTCGCGCCCCAATGCCAATGCTGCTTGGGCATTA 109384

Qy   524  AGTAAGGTTATGCCTTGCCTTTGATTTTTAAATTGATGATTGGTAAAGTTTTTGTTTTTGA 583
Db 109385 AGTAAGGTTATGCACGACCACCTTTTATCTTTTAGTTCATTTTACTGGCGATATAGAGACGCTCTGTTCA 109444

Qy   584  TGATATAATGCCATAACCACCATCACTCAAAATAAATAAAAAATCTGCCGGTTTGGTGGGCTA 643
Db 109445 CGTGCCAAATAAAGTCTGCCCTGCTGCCATCCGTGACT------AGCCC 109485

Qy   644  AGTAAGCTATTTGTTGTGATAAATGGTGTGATTTGGGGCAACAGCTTGTCAGTGTGTCAAG 703
Db 109486 ATAGTGCCCTTGTTAGGAGCATGAATACCAAGTGCTCAGTGCGCGGCAAACTAACTATCTGACAC 109545

Qy   704  CATTGCTTGGCTAAATTGGGCATGCGTCTTGGGTAGATGATGGCGAAATTTCAAAA 763
Db 109546 CAATGGTCGCGGGTCAA---CGCAATCTTCACGATTAATCACTACCGATTGGCCAC 109603

Qy   764  ATACCGCAAATTTTGTCCAAATAATCCACCATGCATCGTGCCCGATCAAGATGATCAGCG 823
Db 109604 ATACGATGTTTGCCTTGGGTGATAGGCGGACGCATATCACCGTGACGGGTCTAAAATGATCAGGC 109663

Qy   824  GAGAGATTTAAAAATTGTTGCCCCTTGAGCTGCCAAAATTGGGAATAGCTCAAGCTGAAAA 883
Db 109664 GACAATATTGAGGACTGTTGGCAAGTCCACCTGGCGCCCTAAGTTGGTTAATAATTCAAAGTAA 109723

Qy   884  CTTGACAGCTCAAGCACCACCAAAATCCATATTCCATACTTTATCATTACTCAAAAGCTCA 943
Db 109724 CTAGATAATTCTATTAACTGCTAGCTCCATTTTATCAATTACTTCAAAAGGCTCAGCGCCTGGC 109783

Qy   944  GTGCCAATAATTACCCCGACACCAACAATCAATGCCTGCANGTTTTTGCCAATCTCGCCTACT 1003
Db 109784 ACGCCAATTATTACCGGCCAACATTGATCACGCCAAACATCGCCGCCCATTGTCCAACC 109843

Qy  1004  AATGTCGTAACAGAGTGCTTTTGGCATTGAGCCGTCTGAGTGGCGGATGATTGG 1053
Db 109844 AGCGGTCGTCACAGTGCTTTTTGCATTTGACCCAGTAATCGCAACAATTGG 109893

Qy     2  AAAGGGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATCC 61
Db 109205 ACAGCCGCGCCATGGTTGAGCCAATATTGGTGCCTTTAGAATCATTAAAGTAATCAAATGCCA 109264
```

RESULT 8
AX932204/c

```
LOCUS       AX932204                1314 bp    DNA     linear   PAT 22-DEC-2003
DEFINITION  Sequence 157 from Patent WO03087353.
ACCESSION   AX932204
VERSION     AX932204.1  GI:40312625
KEYWORDS    
SOURCE      Haemophilus influenzae
  ORGANISM  Haemophilus influenzae
            Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
            Pasteurellaceae; Haemophilus.
REFERENCE   1
  AUTHORS   Edwards,A., Dharamsi,A., Vedadi,M., Domagala,M., Houston,S.,
            Awrey,D., Beattie,B., Mansoury,K., Ouyang,H., Vallee,F.,
            Richards,D., Nethery,K., Virag,C., Buzadzija,K., Pinder,B.,
            Alan,M.Z., Tai,M., Canadien,V. and Kanagarajah,D.
  JOURNAL   Patent: WO 03087353-A 157 23-OCT-2003;
            Affinium Pharmaceuticals, Inc. (CA)
FEATURES             Location/Qualifiers
     source          1..1314
                     /organism="Haemophilus influenzae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:727"
ORIGIN

Query Match         11.6%;  Score 160.4;  DB 6;  Length 1314;
Best Local Similarity   49.6%;  Pred. No. 7.5e-28;
Matches  683;  Conservative   0;  Mismatches  606;  Indels  89;  Gaps  7;
```

```
RESULT 9
U32793/c  U32793
LOCUS               U32793         11498 bp    DNA     linear   BCT 02-JUN-2004
DEFINITION  Haemophilus influenzae Rd KW20 section 108 of 163 of the complete
            genome.
ACCESSION   U32793 L42023
VERSION     U32793.1  GI:1574683
KEYWORDS    .
SOURCE      Haemophilus influenzae Rd KW20
  ORGANISM  Haemophilus influenzae Rd KW20
            Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
            Pasteurellaceae; Haemophilus.
REFERENCE   1  (bases 1 to 11498)
  AUTHORS   Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A.,
            Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J., Dougherty,B.A.,
            Merrick,J.M., McKenney,K., Sutton,G.G., FitzHugh,W., Fields,C.A.,
            Gocayne,J.D., Scott,J.D., Shirley,R., Liu,L.I., Glodek,A.,
            Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E.,
            Cotton,M.D., Utterback,T.R., Hanna,M.C., Nguyen,D.T., Saudek,D.M.,
            Brandon,R.C., Fine,L.D., Fritchman,J.L., Fuhrmann,J.L.,
            Geoghagen,N.S., Gnehm,C.L., McDonald,L.A., Small,K.V., Fraser,C.M.,
            Smith,H.O. and Venter,J.C.
  TITLE     Whole-genome random sequencing and assembly of Haemophilus
            influenzae Rd
  JOURNAL   Science 269 (5223), 496-512 (1995)
  PUBMED    7542800
REFERENCE   2  (bases 1 to 11498)
  AUTHORS   Tatusov,R.L., Mushegian,A.R., Bork,P., Brown,N.P., Hayes,W.S.,
            Borodovsky,M., Rudd,K.E. and Koonin,E.V.
  TITLE     Metabolism and evolution of Haemophilus influenzae deduced from a
            whole-genome comparison with Escherichia coli
  JOURNAL   Curr. Biol. 6 (3), 279-291 (1996)
  PUBMED    8805245
REFERENCE   3  (bases 1 to 11498)
  AUTHORS   White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-SEP-1997) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
  REMARK    The H. influenzae sequence has been updated by R. Fleischmann. New
            database matches have been assigned, product names have been
            improved, and a number of frame shifts have been corrected. We
            gratefully acknowledge the work of Tatusov et. al. We have
            incorporated their annotation into the /notes fields of the
            corresponding H. influenzae genes
REFERENCE   4  (bases 1 to 11498)
  AUTHORS   White,O., Clayton,R.A., Kerlavage,A.R. and Fleischmann,R.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JUL-1995) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
REFERENCE   5  (bases 1 to 11498)
  AUTHORS   White,O., Clayton,R.A., Kerlavage,A.R., Fleischmann,R.D.,
            Peterson,J., Hickey,E., Dodson,R. and Gwinn,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAY-1998) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
  REMARK    The whole genome was shifted by 588 nucleotides for a new start
COMMENT     On Oct 1, 1996 this sequence version replaced gi:1221887.
FEATURES             Location/Qualifiers
     source          1..11498
                     /organism="Haemophilus influenzae Rd KW20"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:71421"
     gene            131..586
                     /gene="HI1129"
     CDS             131..586
                     /gene="HI1129"
                     /note="similar to SP:P22186 GB:X52063 PID:40849 PID:42319
                     GB:U00096 percent identity: 59.2; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAC22784.1"
                     /db_xref="GI:1574684"
                     /translation="MFRGATAVNLDSKGRVAIPTRYRAEILRKNQGQMVCTVDIRQSC
                     LLIYPLDSWEKIEQKLLALSNPDTQRRLQRVWLGHATECEMDAQGRILLSGPLRQHA
                     KLRKGLMLVGQLNKFEIWSDVEWHTQIABDIEIGSSTDFAADALNDFSL"
     gene            617..1582
```

```
gene            617..1582
```

```
     CDS             /gene="HI1130"
     617. .1582
                     /gene="HI1130"
                     /note="similar to GB:D10483 SP:P18595 GB:X52063 PID:216497
                     PID:40850 percent identity: 62.09; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAC22785.1"
                     /db_xref="GI:1574685"
                     /translation="MNSENSFSSSEHITVLLHEAVNGLALKENGIYIDGTFGRGGHSR
                     FILSQLSSNPQLDEAERGFSFMKDGPLDMRDTTQGLSAEEWLKQVSIEDLTWVLKTF
                     ILLDLGVSSPQLIEAERGFSFMKDGPLDMRDTTQGLSAEEWLKQVSIEDLTWVLKTF
                     GEERFAKRIATAIVDFNKSAVKNGTEFLSRTSQLAEIISQAVPFKDKHGHPATRSFQA
                     IRIFINSELDELSLLNSALDMLAPEGRLSIISFHSLEDRMVKHFMKQSKGEDIPKG
                     LPLREDQIQRNQKLRIIGKAIQPSDAEIQANPRSRSAILRVAERI"

     gene            /gene="HI1131"
     1585. .1908
                     /gene="HI1131"
     CDS             1585. .1908
                     /gene="HI1131"
                     /note="similar to SP:P22187 GB:K00137 GB:S49802 GB:S49875
                     GB:X52063 percent identity: 36.63; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="cell division protein (ftsL)"
                     /protein_id="AAC22786.1"
                     /db_xref="GI:1574686"
                     /translation="MSENNKPRYPLQQIIVEDLFSSNKLVLVLLLIGILVSAMGTIWIT
                     HKTRQLISENGMLILQRQALENEYRNLQVQEATEGDSTRVESIAISTLKMKVVSSEQE
                     VEIRE"

     gene            /gene="HI1132"
     1921. .3753
                     /gene="HI1132"
     CDS             1921. .3753
                     /gene="HI1132"
                     /note="similar to SP:P04286 GB:K00137 GB:S49802 GB:S49875
                     GB:S49875 percent identity: 53.11; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="penicillin-binding protein 3 (ftsI)"
                     /protein_id="AAC22787.1"
                     /db_xref="GI:1574687"
                     /translation="MVKFNSSRKSGKSKKTIRKLTAPETVKQNKPQKVFEKCFMRGRY
                     MLSTVLILLGLCALVARAAYVQSINADTLSNEADKRSLRDKVRSSILDRNGQLL
                     SVSVPMSAIVADPKTMLKENSLADKERIAALAEELGWTENDLVKKIEKNSKSGYLYLA
                     RQVELSKANYIRRLIKGILLETEHRRRYPRVEEAAHVGYTDIDGNGIEGIEKSFNS
                     LLVGKDGSRTVRDKDRGNIVAHISDEKKYDAQOVTLSIDEKLQOSMVRFRIKAVSENN
                     AESGTAVLVDVERTGEVLAMATAPSYNPNNRVGVKSELMRNRAITDTFEPGSTVKPFV
                     ITALQRGVKRDEIIDTTSFKLSGKEIVDVAPRAQQTLDEILMNSSNRGVSRLALRMP
                     PSALMETYQNAGLSKPTDLGLIGEQVILINANRKRWALRERATVAYCIGYTATPLQIA
                     RAYATLGSFGVVRPLSITKVDPPVIGKRVFSEKITKDIVGILEKVAIKNRAMVEGYR
                     VGVKTGTARKIENGHYVNKYVAFTAGIAPISDPRYALVVLINDPKAGEYYGGAVSAPV
                     FSNIMGYALRANAIPQDAEAAENTTTKSAKRIVYIGEHKNQKVN"

     gene            /gene="HI1133"
     3763. .5229
                     /gene="HI1133"
     CDS             3763. .5229
                     /gene="HI1133"
                     /note="similar to GB:D10483 SP:P22188 GB:X55814 PID:285768
                     PID:581032 percent identity: 55.72; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="UDP-N-acetylmuramyl-tripeptide synthetase
                     (murE)"
                     /protein_id="AAC22788.1"
                     /db_xref="GI:1574688"
                     /translation="MKKLTALFNLPELKNDIELHLHNMVLDSRKYKAGDLFVAIKGHQVD
                     GNQFIDSALHSGASAAVVSETELSLSSEHLTVAFIGNVPVVKKYQYLQAHHLSSLADVYDSP
                     SNNLLTVGTVGTIGKTTISOLLAQWELLGHRAAVMGTIGNGLFGQIVEAKNTTGSAV
                     EIQSLSAFKHAGADFTSIEVSSHGLAQHRVEALHFKAAIFTNLTRDHLDYHOSMENY
                     AAAKRLFTELDTQIKVINADDEIGYQWLTELPDAIAVSMNADFKVGSHOWMKAINIH"

     gene            /gene="HI1134"
     5243. .6616
                     /gene="HI1134"
     CDS             5243. .6616
                     /gene="HI1134"
                     /note="similar to GB:D10483 SP:P11880 PID:216500 PID:40854
                     PID:42048 percent identity: 51.88; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="UDP-MurNac-pentapeptide synthetase (murF)"
                     /protein_id="AAC22789.1"
                     /db_xref="GI:1574689"
                     /translation="MIKLSTVQLAQILQAKLIGDENVQVEKINTDTRKSVSNSLFFAL
                     KGEKFDAHQILDQAVSQGALALVVQQENSSISVPGLVVKDTRIALGELAKWLREKINP
                     RTVAMTGSSGKTTVKEMTASILCHGTAADSEAVLFTNGNNFNNDIGVPLTLLRTHRHF
                     AVIELGANHQNEINYTTKLVQPNAALINNIAPAHLEGFGSLAGVVQAKGEIYRGLTKN
                     GVAIINAEHNHLDIWQKEISNHAIQYFNGKDYSAKNIHHTSQGSTFTLISPQGEIEIT
                     LPYLGEHNVKNALAAIALAAMVGAILTDVKAGIEQRSQVKRGRLPIQVTPNLLLLDDT
                     YNANKDSLCAAIDVLKGYDAFRIICVGDMKELGENSLAIHREVQGYINLVNLDLVCSY
                     GNESAVISEAVSGKHFTDKTEMVDFLVPLIENQLQQNKVVVLGKGSRSMKMEDVIYS
                     LKDKIKC"

     gene            /gene="HI1135"
     6610. .7692
                     /gene="HI1135"
     CDS             6610. .7692
                     /gene="HI1135"
                     /note="similar to GB:D10483 SP:P15876 GB:X51584 PID:216501
                     PID:40855 percent identity: 76.67; identified by sequence
                     similarity; putative"
                     /codon_start=1
                     /transl_table=11
                     /product="phospho-N-acetylmuramoyl-pentapeptide-
                     transferase E (mraY)"
                     /protein_id="AAC22790.1"
                     /db_xref="GI:1574690"
                     /translation="MLVWLAEYLVRYETAFNAISYITVRANLALLTALFISLWIGPKV
                     IKRLQILKPGQEVRNDGPESHFAKKGTPTMGGVNMILFSIGVSTLLWANLANPYIWVCL
                     FVLFGYGAIGFVDDFRKITRKNTDGLIARWKYFWMSVVALVAVIAILMLYWLGHDTDATRL
                     VIPFKDIMPQLGLPFYIVLSYFVIVGTGNAVNLTDGLDGLAIMPTALVAGAFALIAWA
                     TGNVNFAEYLHIPYIKYSSEVVVFCTAIVGAGLPFWFNTYPAQVFMGDVGSLALGGA
                     LGVVAILVRQEFLLVIMGGVFVVEALSVILQVGSYKLKRQRIFRMAPIHHFELKGWP
                     EPRVIIRFWIISLMLVLMGLVTLKLR"

     gene            /gene="HI1136"
     7815. .9128
                     /gene="HI1136"
     CDS             7815. .9128
                     /gene="HI1136"
                     /note="similar to GB:D10483 SP:P14900 SP:P20100 GB:M30807
                     GB:X17609 percent identity: 60.96; identified by sequence
                     similarity; putative"
```

```
Query Match          11.6%;    Score 160.4;   DB 1;   Length 11498;
Best Local Similarity    49.6%;   Pred. No. 5.6e-28;
Matches   683;   Conservative    0;   Mismatches   606;   Indels   89;   Gaps   7;

Qy        1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCAATTATAATTTTTAAATTGATC      60
            ||||   ||||| ||||||| |||||| |||||||||||  ||||||| ||| |||||||||
Db     9122 TAAACATTGAGCTAAATGCGTAAAATCTTCGCCGCGCTTTCAAAAAGAAGCAAAACTGATC    9063

Qy       61 CATGCTTAGCACCAACAGCAGGTGACAAAAGCACAGCCCTAGCGCGTGCTTTGACTGGCTGGCCAA     120
            ||||| ||| ||||||||| |||| |||| ||||| |||||  ||||||||||||||| ||||||||
Db     9062 GAGACTTGCACAAGCAGGCGCGACAATAATACCA-----------TATCTCCGCTTGCAA    9015

Qy      121 TACTGCCCGATCAAGTGTACCAAACCTGATATATATTTAGATGCCAAATGTGCATCACCTTT     180
            ||||  ||| || |||| ||| ||||||| |||||||  || |||||||| |||||||||||
Db     9014 TGTTGGGCGTAAAAATTCTATACGCTTGTTCCCATTGTATCGAACAA--------ATAACTTTG    8961

Qy      181 GAGTAAATCTTGCTCAATTATCCGGCGGCCATCTTGACCGGATTAAATACCACACTATCCACCCA     240
            ||||||||||  |||  |||||||| ||| ||||| |||| ||||||| ||| || ||||||||
Db     8960 CGATGAAAATTTGCAAGCACCGCACCCATCTCGACCAAAAACAATAACAAAATAAGTGCTGTGG    8901

Qy      241 CTGCATAACAATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCGCCACCCAAAAT     300
            |||   ||||   |||||| |||| |||||||||||||||  ||||||| |||||||||||||
Db      241 ...
```

RESULT 10

LOCUS   AX191762   11498 bp   DNA   linear   PAT 15-AUG-2001
DEFINITION Sequence 44 from Patent WO0149775.
ACCESSION   AX191762
VERSION   AX191762.1 GI:15209931
KEYWORDS   .
SOURCE   Haemophilus influenzae
  ORGANISM Haemophilus influenzae
            Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
            Pasteurellaceae; Haemophilus.
REFERENCE   1
  AUTHORS   Iversen,P.L.
  TITLE   Antisense antibacterial cell division composition and method
  JOURNAL   Patent: WO 0149775-A 44 12-JUL-2001;
            Avi Biopharma, Inc. (US)
FEATURES   Location/Qualifiers
     source   1..11498
                     /organism="Haemophilus influenzae"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:727"
ORIGIN

Query Match   11.6%; Score 160.4; DB 6; Length 11498;
Best Local Similarity   49.6%; Pred. No. 5.6e-28;
Matches 683; Conservative   0; Mismatches 606; Indels   89; Gaps   7;

Qy   1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATATTTTTAAATTGATC   60
               ||| |||| | || ||||| || | |||| | | || || | ||| || |||| ||||
Db   9122 TAAACATTGAGCTAAATGCGTAAAATCTTCGCCGGCCTTTTCAAAGAAGCAAAACTGATC 9063

Qy   61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTTAGCGGGTGCTTTGACTGGCTGCCAA   120
               |||||| | ||||| ||| | || ||||| |||||| |||| |||| |||| ||||
Db   9062 GAGACTTGCACAAGCAGGCGGACAAGAATAATACCA----------TATCTCCGCTTGCAA 9015

Qy   121 TACTGCCCGGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAGTGCATCACCTTT   180
               |||| | | || | || || | | | || | |||| | | || ||| ||
Db   9014 TGTTGGGCGTGAAAAATTCTATCGCTTGTTCCATTGTATCGGAACAA-------ATAACTTTG 8961

Qy   181 GAGTAAAATCTTGCTCAAATTATCGGGGCCATCTTGACCGGATTAAAATACACACTATCCACCCA   240
               || ||||||| | | | | || | || | |||| | | | || || | ||||
Db   8960 CGATGAAAAATTTTGCAAGCAGCGCAACCATCGTTGACCAAAATAACAAATAAAGTGTGG 8901

Qy   241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAT   300
               ||||| | | | ||||| | |||||| | || | | | ||| || ||| |||
Db   8900 TTGATTAATTAATTCAGCGAATTCTGAAAAAATCAGCCCTTTTCCGTCTCCGCCTAGCAA 8841

Qy   301 CAATGCCAGTGATGATATGCCATAAAACCGTGCCAAGCCCATCAATTGCTGCCAATGGTTGA   360
               ||||| ||| ||| | | || ||||| || | || |||||||| | |||| ||| | |||
Db   8840 CAAATGCAATTACCCTCAATATA------AAGCCCAGCCAATGCAGCAACTGTTACT 8790

Qy   361 GCCCATGTTGTTACCTTTTGAATCATTATAAAAATATCATCACTAATCATCACAATATA   420
               |||||| | ||| ||| || | || | ||||||| | || || | | || ||
Db   8789 CCCCACATTTGTTGCTTTAGAGTCAATTAATCCAACGAATGCCATTAGCTTGATGCACTAA 8730

Qy   421 TTCACAGCGATGCGGTAGCCCTTTGAAGTCTTAAGGGTATCAAGCATAGGGCATCCATGGG   480
               ||||| || || || | || | |||| |||| | | |||| || || |||| ||||
Db   8729 TTGAAAACGATGCTAACCCTTTGAAAATGACGAAGTGCGTACCAATTGAAATCTAAAATT 8670

Qy   481 AAGATTGGGCAGCTGTGCCTAAGGCAAGTGGCAGATAAGGCATTGAGTAGGGTTATGCTTGCC   540
               |||| || | | || || ||||||||| || |||||||| || | | ||| |||| |||
Db   8669 AATACCTATAGCTTGTGCCAATGCTGTTGGCTGCCAAAAATGTTCATATAATTATGGCGACC 8610

Qy   541 TTTGATTTTTAATTGATGGATTGGTTAAAAGTTTTGTTTTTGATGATATAAATGCCATACC   600
               ||||||||| | || ||| || | |||| ||| |||||||||||||| || ||||||||
Db   8609 AACCAAATGTAGCTTCTTCACAAGGTAAAATCACTTCATCTTTTACCATTAAAATATTGCTT 8550

Qy   601 ATCACCATCATCAAAATAAAATAAAAATCTGC-CGTTTGGTGTGGGCTAAGTAAGCTATTTGTTG   659
               |||| | || || || ||| | ||||| || |||| || | || ||| || |||| |||
Db   8609 ATCACCATCATCAAAATAAAATAAAAATCTGC-CGTTTGGTGTGGGCTAAGTAAGCTATTTGTTG 659

Qy   660 TGATAATGGTGTGATTGGGGGCAAGCGCTTGTCAGTGTGGTCAAGCCATTGCTTGGCTAAAT   719
               || ||| || || ||| ||||| || ||| |||| || ||||| |||| || || ||||
Db   8669 AATACCTATAGCTTGTGCCAATGCTGTTGGCTGCCAAAAATGTTC------ 8610

Qy   720 TGGCATCGTCTTGGTAGATGATGGCGGGTTGGCAAATTTCAAAAATACCGCAATTTTGTG   779
               |||| || | ||| || | ||| | ||| |||| |||| || ||| | || ||||| |
Db   8457 ---TATCTTCATTGTTCAACACACTACTTAGCATTAATGATAAATATTTTTAAAATTTTTGTG 8401

Qy   780 CCAAATAATCCACACAGCATCCGTCCGATCAAGATGATCAGCGGGAGAGATTTAAAATTG   839
               |||| ||||| || ||| |||| || |||| || |||| | | ||| |||| ||| ||||
Db   8400 GGCCATAATCTTCTAAAATCCATAGCGATCCAATGACCTTCAGTCACGTTCAAGACAG 8341

Qy   840 TTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAACTTGACAGCTCAAGCA   899
               ||||||||| || |||||||||| || | || |||| |||| | |||| |||| |||||
Db   8340 TCCGTGCCGCAGCTTTTAAGCTATAAGGTTGTCTCAAGCTGAAAAGCTCTAGTA 8281

Qy   900 CCACCAAAATCCATATTTTGGATCGTTAAGTGGGTGCCAATATTACCGC   959
               |||| ||| | | | ||| || | |||| || | |||| || ||||
Db   8280 CATAAAGTTCACAAATCTTCATTC---AACAATGACAAAGCGGAATCCCAATATTTCCGC 8224

Qy   960 CGACACCAACAATCATGCCTGCATCTCGCATGTTTTGCCAATCTCGTACTACTAGTAGTGC   1019
               | |||| ||| || || | | ||||| | |||| ||||| | | | | | || |
Db   8223 CCATACCAACTTTCACACAGCTTTCGGCCAGTTTGTGTTGTGTTCATAAACTAAGTACGGTAC 8164

Qy   1020 TTTTGGCATTGAGCCCTGTGATGGGCGGATGATGATTGGGTGTGTTGCTTGAGTTAGATCTC   1079
               |||||| ||||| || ||||| |||| || ||| || || | ||||| ||| ||| |||
Db   8223 CCATACCAACTTTCACACAGCTTTCGGCCAGTTTGTGTTGTGTTCATAAACTAAGTACGGTAC 1079

Qy   1080 GTGCTTTGAGTGTATCAACAAATGAACAAATTGAAATCGATTACTTCCACTCCCGCTTAA   1139
               ||||||||| || || ||| ||| | ||||| | | | | ||| ||||| |||| ||
Db   8163 TTTTACCATTGAACCTGTAATCCCAAATTGTTGTCGCTTGTGTTGTTCGCGTTGAGTTAGATCTC 8112

Qy   1140 TGGCCGGGGACAACTACTCCGCGTTCTTCGTGGGTTAATGCCGGGGCTGATGATGATCTCGG   1199
               | | || ||| | || |||| || ||||| ||||| | | ||| || ||| | |||| |
Db   8111 ------GAATAATTCAAATATCGCGACTGCGTACTTCCACTGCGTCCTTTAA 8071

Qy   1200 CTGAGCTTAAGAGTTCACTGTCCAAAAATGACAAATGAACGGCCTGGTGGGCAGTT   1259
               |||| ||||||| | || |||| || | |||||| || || |||| | |||||||
Db   8070 GTCCGGTTTGAATTTCTGGTGTTTTTACGGCAAGCCCTGTAATACAAATCATATCGC 8011

Qy   1260 TTTCAAGTAACCATTCCTGATTTAAAACTACCAGTATAATTTCTGTTCGTTGACCGTGATTGA   1319
               ||| ||||| |||| | ||||| ||| ||| ||| | |||||| | || | ||| ||| |
Db   8010 TTTCAGTAACCATTCCTGATTTAAAACTACCAGTATAATTTCTGTTGTTGACCGTGATTGA 7951

Qy   1320 CCAAAAAATTCACAGCTGACAGTCGCCCAAACCGACGACGCGCCATATTTTTT   1377
               ||||||| ||||| | ||| ||| | | ||| | || | |||||||| ||
Db   7950 TATCAATACCAGTAGGAGATTTTTCGGAGTGATCAATCAACAGAAATTAGCCTGTTTGGGGATA 7891

Db   7890 AGAGTAAATCCACACAGAAGAAAGACCTGTTTTGCCAAGCCCGATGATAGAGTAATATTTTT 7833

```
Qy        660 TGATAATGGGTGTGATTGGGGGCAACGGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAAT 719
              ||||        ||      |||
Db       8489 CGCTTGATTTTTCGTTTTTCCCAAAAGTCAGCC------------------------- 8458

Qy        720 TGGCCATCGTCTTGGTAGATGAGTGGCGGTTGGCAATTTCAAAAATACGCAATTTTTGTG 779
              |||||
Db       8457 ---TATCTTCATTGTTCAACACACTACTTTAGCATTATGATAAATTGCGTAATTTTTGCTT 8401

Qy        780 CCAAATAATCCACCATGCCATCGTCCGATCAAGATGATCAGCGGAGAGATTAAAATTG 839
                            |||
Db       8400 GGCGGATAATCTTCTAAATCCATATAGCCGATCCATATGAACTTCAGTGCACGGTTGCAA 8341

Qy        840 TTGCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTCAAGCTGAAAAAACTTGACAGCTCAAGCA 899
              |||||
Db       8340 TCGCTGCCCGCAGCCTTTTAAGCTTATAAGTTGTCTCAAGCTGAAAAACTAGAAAGCTCTAGTA 8281

Qy        900 CCACCAAATCCATATATTTGGATCGTTAATAATCAAGTGACAGTGTCCAATATTACCGC 959
              |||
Db       8280 CATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAATATTACCGC 8224

Qy        960 CGACACCAACAACAATCATGCCTGCATGTTTTCCCATCTCGCTCTACTAATGTCGTAACAGTGC 1019
              |||
Db       8223 CCATACCAACTTTCACACCAGCACTTCGCCATTTCATAAACTAAAGTAGTTACGGTAC 8164

Qy       1020 TTTTTGGCATTTGAGCCTGTGATGGGACGTGATGATTGGTGTTGTTGTTTGCTTGAGTTAGATCTC 1079
              |||||
Db       8163 TTTTTACCATTTGAACCTGTAATCCCCACAATTGGCTTTGTCGCTCGCGGCA--------- 8112

Qy       1080 GTGCTTGAGTGTATCAAGTGAACAATTGAACATGCTGATCACGACGAATGCCTTGAGCTT 1139
              |||
Db       8111 ---------------GAATAATTCAAATATCGGCGATTACTTCCACTCCCGCTTTAA 8071

Qy       1140 TGGGCGGCGACAAATACTCGGCGCGTTCTTGGGTGTTTTTACCGCAAGCCCTGGTCGTTATCGG 1199
              |||||
Db       8070 GTGTCGGCGTTTGAATTTTCTGGTGTTTTTACCGCAAGCCCTGGTCGTTATCGG 8011

Qy       1200 CTGAGCTTAAAGAGTTCACTCGTCCAAGCTGCCAAAATGACAATGAAACGGCTTGGGGCAGTT 1259
              |||
Db       8010 TTTCAAGTAACCAATCCTCGATTTAAAACTACCAGTTGAGAGGGATATTTTGAGGGAAGTT 7951

Qy       1260 TATCTGCCAAGGTAGGATGGGAATTACCATCGATGAGATGCTAACTTTGTGACCGTGATTGA 1319
              |||
Db       8007 TATCAATACCAGTTAGGATTTTTTCGGTAATCAATCACACAGAAATATTAGCCTGTTCGGGGATA 7891

Qy       1320 CCAAAAAATTCACAGCTGACAGTCGCCAAAACCGACCACGAACATATTAGCCCGATGATGTAATATATTTTTT 1377
              |||||
Db       7890 AGAGTAAATCCACACACAAGAAAAGACCTGTTTTGCCAAGCCCGATGATGAGTAATATATTTTTT 7833
```

```
Qy          1 TAAAGCCTGAACCATGTTAACAAACTTATCACCTCGCCTCATTATAATTTTAAATTGATC 60
              ||||
Db     104439 TAAACATTGAGCTAAAATGCGTAAAATTCTTCGCGCGGCGTTTTCAAAAGAAGCAACTGATC 104380

Qy         61 CATGCTAGCACAAGCAGGTGACAATAAGCACACAGCCCTAGGGGTGCTTTGACTGGCTGCAA 120
              |||
Db     104379 GAGACTTGCACAAGCAGGCAGGCGGACAATAATACCA------TATCTCCGCTTTGCAA 104332

Qy        121 TACTGCCCGATCAAGGTGTACCAAACCTCGATATATTTTTAGATGCGCATCACCTTTT 180
              |||
Db     104331 TGTTGGGCGTAAAAATTCTATCGCTTGTTCCATTGTATCGAACAA--- ATAACTTTG 104278

Qy        181 GAGTAAAATCTTGCTCAATTATCGGGGGCCATCTTGACCGGATTAAATACACTAACTAATGTGG 240
              |||
Db     104277 CGATGAAAATTTTGCAAGCACGAGCGACCACATTCCGACCAAAACAATAACAAATAATGTTGG 104218

Qy        241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGAAAGTTCAAGCATTGCAGACTAAAAT 300
              |||
Db     104217 TTGATTAATTAATTCAGCTAATTCTGAAAAATTCAGCCCTTTTTTCCGTCTTCCGCCTAGCAA 104158

Qy        301 CAAATGCCAGTGATGATATGCCATAAAACCGTGCCAAGCCCATCAATTGCTCGTCGAAATGGTTGA 360
              |||
Db     104157 CAAATGCAATTTACCCCTCAAATA---------AAGCCCAGCGAATGCAGCAAACTGCTT 104107

Qy        361 GCCGATGTTTTGGTTTACTTTTTGGAATCATTTAAAAATATGGCCTTGCTATCAATATCAATATA 420
              |||
Db     104106 CCCCACATTTGTTTCGTAGAGTCATTACTAGTGATAATAGTCAGAAAT 104047

Qy        421 TTCACAGCGATGCGGTAGCCCTTTGGCAATTTCAAAAATACGCAATTTTTGCTT 480
              |||
Db     104046 TTGAAAACGAGATCTAACCCTTTGAAATGACGAAGTTGCGGGTACGAATTGAAATCTAAAT 103987

Qy        481 AAGATTGGCAGCTGCGTGCCCTAAGGCAAGTGCAGATAAGGCCAATTGAGGTTTATTGCTTGCC 540
              |||
Db     103986 AATACCTATACGCTTCTTCACAAGGTAAATCACTTTCCATCTTTTTTTTGGATGAAAATTAATGCCGAC 103927

Qy        541 TTTGATTTTTTAATTGATGGATTGGTAAAAGTTGTTTTTGTGATGATATAATAATTGCTT 600
              |||
Db     103926 AACCAAATGTAGCTCTTCTCACAAGGTAAAATCACTTCATCTTTACCATTAAAATATGCTT 103867

Qy        601 ATCACCATCATCAAATAAATAAAAATCTGC--CGTTTGGGTGGCTAACTTAAGCTATTTGTTG 659
              |||
Db     103866 GCCAATTTTCAGTTTTTTAGCCAATAATCCGACGACTATTTTTCGCGTGTGTCAAGCATTGCTTGGCTAAAT 103807

Qy        660 TGATAATGGGTGTGATTGGGGGCAACGGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAAT 719
              |||
Db     103806 CGCTTGATTTTTCGTTTTTCCCAAAAGTCAGCC-------------------------- 103775

Qy        720 TGGCCATCGTCTTGGTAGATGAGTGGCGGTTGGCAATTTCAAAAATACGCAATTTTTGTG 779
              |||
Db     103774 ---TATCTTCATTGTTCAACACACTACTTTAGCATTATGATAAATTGCGTAATTTTTGCTT 103718

Qy        780 CCAAATAATCCACCATGCCATCGTCCGATCAAGATGATCAGCGGAGAGATTAAAATTG 839
              |||
Db     103717 GGCGGATAATCTTCTAAATCCATATAGCCGATCCATATGAACTTCAGTGCACGGTTGCAA 103658

Qy        840 TTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA 899
              |||
Db     103657 TCGCTGCCCGCAGCCTTTTAAGCTTATAAGTTGTCTCAAGCTGAAAAACTCTAGTA 103598

Qy        900 CCACCAAATCCATATATTTGGAATCGTTAATAATCAAGTGACAGTGTCGCAATATTACCGC 959
              |||
Db     103597 CATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAATATTACCGC 103541

Qy        960 CGACACCAACAACAATCATGCCTGCATGTTTTCCCATCTCGCTACTTAATGTCGTAACAGTGC 1019
              |||||
```

RESULT 12
WPCOMMENT
BD426631_12/c
Sequence split into 19 fragments   LOCUS BD426631 Accession BD426631

Fragment Name	Begin	End
BD426631_00	1	110000
BD426631_01	100001	210000
BD426631_02	200001	310000
BD426631_03	300001	410000
BD426631_04	400001	510000
BD426631_05	500001	610000
BD426631_06	600001	710000
BD426631_07	700001	810000
BD426631_08	800001	910000
BD426631_09	900001	1010000
BD426631_10	1000001	1110000
BD426631_11	1100001	1210000
BD426631_12	1200001	1310000
BD426631_13	1300001	1410000
BD426631_14	1400001	1510000
BD426631_15	1500001	1610000
BD426631_16	1600001	1710000
BD426631_17	1700001	1810000
BD426631_18	1800001	1830121

Continuation (13 of 19) of BD426631 from base 1200001 (BD426631 The Nucleotide Sequence

Query Match   11.6%;   Score 160.4;   DB 6;   Length 110000;
Best Local Similarity 49.6%;   Pred. No. 4.2e-28;
Matches 683;   Conservative 0;   Mismatches 606;   Indels 89;   Gaps 7;

Qy   1 TAAAGCGTGAACCATGTTAACAAGTTATCACCTCGCTCATTATATAATTTTAAATTGATC 60
Db   4439 TAAACATTGACAGCTAAAAGCACAGCCCTAGGGGTGCTTTCAAAAGAGCAAAACTGATC 4380

Qy   61 GAGACTTGCACAAGCAGGGGACAATAATACCA-----------TATCTCCGCTTTGCAA 120
Db   4379 ...

Qy   121 TACTGCCCGATCAAGTGTACCAACCTGATATATATTTTAGATGCCAAAATGTGCATCACCTTT 180
Db   4331 TGTTGGGGGCTAAAAATTCTATCCGTCTTGTTCCATTGTATCGAACCA 4278

Qy   181 GAGTAAATCTTGCTCAATTATCGGGGCATCTCGACCTGATTAAATACAACTATCCACCCA 240
Db   4277 CGATGAAAATTTTGCAACAGCGCCACCATCTCGACCAAAACAATAACAAATAATGTGTGG 4218

RESULT 13
AR274513-11/c
WPCOMMENT
Sequence split into 19 fragments   LOCUS AR274513 Accession AR274513

Fragment Name	Begin	End
AR274513_00	1	110000
AR274513_01	100001	210000
AR274513_02	200001	310000
AR274513_03	300001	410000
AR274513_04	400001	510000
AR274513_05	500001	610000
AR274513_06	600001	710000
AR274513_07	700001	810000
AR274513_08	800001	910000
AR274513_09	900001	1010000
AR274513_10	1000001	1110000
AR274513_11	1100001	1210000
AR274513_12	1200001	1310000
AR274513_13	1300001	1410000
AR274513_14	1400001	1510000
AR274513_15	1500001	1610000
AR274513_16	1600001	1710000
AR274513_17	1700001	1810000
AR274513_18	1800001	1830121

Continuation (12 of 19) of AR274513 Sequence 1 from patent u

Query Match   11.6%;   Score 160.4;   DB 6;   Length 110000;
Best Local Similarity   49.6%;   Pred. No. 4.2e-28;
Matches 683;   Conservative 0;   Mismatches 606;   Indels 89;   Gaps 7;

Qy   1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATATATTTTAAATTGATC 60
Db   104439 TAAACATTGAGCTAAATGCCTAAATTCTTCGCCCGGCTTTTCAAAAGAGAGCAAACTGATC 104380

Qy   61 CATGCCTAGCACAAGCAGGTGACAAAAGCACAGCCCCTAGCGGTGGGTCTTTGACTGGGTTGCCAA 120
Db   104379 GAGACTTGCACAAGCAGGGACAATAATAATAATACCA 104332

Qy   121 TACTGCCCGATCAAGTGTACCAACCTGATAATATATTTTAGAIITGGCAAAATGTGCATCACCTTT 180
Db   104331 TGTTGGGGCGGTAAAAAATTCTATCGCTTGTTCCATTGTATCCGAACAA 104278

Qy   181 GAGTAAAATCTTGCTTCAAATTATCGGGGCCATCTTGACGGCGATTAAAATACACTATCCACCCA 240
Db   104277 CGATGAAAAATTTTGCAAGCAGCGCACCATCTCGACCAACAAAACAATAACAAATAAGTGTGTGG 104218

Qy   241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACTTTACCCAAGCCCACCCAAAAT 300
Db   104217 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCCTTTTCCGGTCTCCGCCTAGCAA 104158

Qy   301 CAATGCCACTGATGAATATGCCCATAAAACCGTGCCAAGCCCATCAAATTGCTGCAATGGTTGA 360
Db   104157 CAAATGCAATTACCCTCAATATA---------AAGCCCAGCCAATGCAGCCAACTGTACT 104107

Qy   361 GCCCGATGGTTGTACCTTTTGAAATCATTAAAAATATGCCTTGCTATCATCAATATA 420
Db   104106 CCCCACATTGTTCCTTTAGAGTCATTAATCCACGAATGCCATTAGCTTGATGCGACTGGG 104047

Qy   421 TTCACAGCGGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGGCATCCATGTTG 480
Db   104046 TTGAAAACGGATGAATCTAACCCTTTGAAATGACGAAGTGGGGGTACGAAATTGAAATCTAAATT 103987

Qy   481 AAGATTTGGCAGCTCTGCCTAAGGCAAGTGCCAGATAAGGCATTGAGTAGGTTATCTTGCTTGCC 540
Db   103986 AATACCTATAGCTTGTGCTGCCAAAAATGTTCATTATAATTATGCGACC 103927

Qy   541 TTTGATTTTTAATTGATGGGATTGGTAAAAGTTTGTGTTTTTTTTTTGATGATAAATGCCATACC 600
Db   103926 AACCAATGTAGCTTCTTCACAAGGTAAAAATCACTCATCTTTTACCATTAAAATATTGCTT 103867

RESULT 14
AR274513-12/c
WPCOMMENT
Sequence split into 19 fragments   LOCUS AR274513 Accession AR274513

Fragment Name	Begin	End
AR274513_00	1	110000
AR274513_01	100001	210000
AR274513_02	200001	310000
AR274513_03	300001	410000
AR274513_04	400001	510000
AR274513_05	500001	610000
AR274513_06	600001	710000
AR274513_07	700001	810000
AR274513_08	800001	910000
AR274513_09	900001	1010000
AR274513_10	1000001	1110000
AR274513_11	1100001	1210000
AR274513_12	1200001	1310000
AR274513_13	1300001	1410000

Qy   601 ATCACCATCATCAAATAAATAAAAAATCTGC-CGTTTGGTGGTGGCTTAAGTAAGCTATTTGTTG 659
Db   103866 GCCATTTTCAGTGTTGTTTTAGCCAATAATCCGACTACTATTTTCCGCAAAAGAAGCAACGGTATGTTTT 103807

Qy   660 TGATAATGGTGTGATTTGGGGGCGAACGCCGTTGTCGAGTGTCAAGCTCAAAAAGCTCTAAAT 719
Db   103806 CGCTTGAATTCGTTTTCGTTTTGTTTCCCAAAAGTCAGCC------------------------- 103775

Qy   720 TGGCATCGTCTTCTTGGGTAGATGATGGGCGGTTTGGCCAATTTTCAAAAATACGCAATTTTTGTG 779
Db   103774 ---TATCTTCATTGTTCAACACACTACTTTAGCATTATGATAAAATGCGTAAITTTTCGCTT 103718

Qy   780 CCAAAATAATCCACCATCGTGCCGATCAAGAACATGAATCAGCGGGAGGAGATTAAAAATTG 839
Db   103717 GGCGGATAAATCTTCTTAAAATCCATATAGCGATCCATATAGTAACTTCAGTCACGTTCAAGACAG 103658

Qy   840 TTGCCCCTTGAGCTGCCAAAATTGTGAGATATGCTGAAAAACTTGACACGTTCAAGCA 899
Db   103657 TCGCTCGCGCGAGCCTTGTCGTATTAAGCTATGTTCGAAGCTGAAAAAGCTCTAGTA 103598

Qy   900 CCACCAAAATCCATATATTTTGGATCGTTAATAATTCAAAGTGTCAGGCGTGCCAAATGTTACCGC 959
Db   103597 CATAAAGTTCACAAATCTTCATTC--AACAAATGACAAAGCGCGGAATCCCAAATATTTCCGC 103541

Qy   960 CGACACCAACAAATCATGCCTGCGTGTTTCCGCATCTCGCCTACTAATGCGTTGTAACAGTGC 1019
Db   103540 CCATACCAACTTTCACACCAGCACGCTTTCCCATTTCATAAAACTAAAGTAGTTAACGGTAC 103481

Qy   1020 TTTTTGGCATTTGAGCCTGTGATGCGGCGATGACGTGTGTTTGTTGCTTGAGTTGAGATCTC 1079
Db   103480 TTTTTACCATTTGAAACCTGTAAATCCCCACAATCCCGACAATTGTCGCGTCGCGCCGGCA---- 103429

Qy   1080 GTGCTTTGAGTGTATCAAATGAACAAATTGAACATCGCTGACTTGATCACAGAGAAATGCCTTGAGCTT 1139
Db   103428 ------------------GAATAAATTCAAATATCGCGATTCACTCCGCCGCTTAA 103388

Qy   1140 TGGCGGCGACAATACTCGGCGTTCTTGGGTTTGGGTTAATGGCCGGGGGGGCTGAATGATCGTCGG 1199
Db   103387 GTGCGCGTTTGAAATTTCTGGTGTTTTTACCCGCAAGCGGTAATACAACAATCATATCGC 103328

Qy   1200 CTGAGCTTAAGAGTTCACTGTCCCAAGCTGCCAAAAATGACAAATGGACGCCCTGGGCAGTT 1259
Db   103327 TTTTCAAGTAACCATTCCTGATTTAAAACTACCAGTAACGGATATTTTTGAGGAAGTT 103268

Qy   1260 TATCTGCCAAGGTAGGAIITGGGATTATTAACGATGAGTGCTAAACTTTGTGTTTGGGATTGA 1319
Db   103267 TATCAATACCAGTAGGATTTTTTTCGAGTCATCAATCACACGAATAATTAGCCTGTTGGGATA 103208

Qy   1320 CCAAAAAATTCACAGCTGACAGTGCCCAAAACGACGACGGGCATATTTTTT 1377
Db   103207 AGAGATAATCCACACAAGAAAAGACCTGTTTTGCCAAGCCCGATGAGTAGTAATATATTTTTT 103150

```
Db   AR274513_14   1400001
Db   AR274513_15   1500001
Db   AR274513_16   1600001
Db   AR274513_17   1700001
Db   AR274513_17   1810000
Db   AR274513_18   1800001
                   1830121
Continuation (13 of 19) of AR274513 Sequence 1 from patent U
Query Match      683;  Conservative     0;  Mismatches  606;  Indels  89;  Gaps  7;
Best Local Similarity   49.6%;  Pred. No. 4.2e-28;
Matches  683;
```

RESULT 15
AR632719_11/c
WPCOMMENT

Sequence split into 19 fragments LOCUS AR632719 Accession AR632719

```
Fragment Name    Begin        End
AR632719_00      1            110000
AR632719_01      100001       210000
AR632719_02      200001       310000
AR632719_03      300001       410000
AR632719_04      400001       510000
AR632719_05      500001       610000
AR632719_06      600001       710000
AR632719_07      700001       810000
AR632719_08      800001       910000
AR632719_09      900001       1010000
AR632719_10      1000001      1110000
AR632719_11      1100001      1210000
AR632719_12      1200001      1310000
AR632719_13      1300001      1410000
AR632719_14      1400001      1510000
AR632719_15      1500001      1610000
AR632719_16      1600001      1710000
AR632719_17      1700001      1810000
AR632719_18      1800001      1810000
                 1830121
Continuation (12 of 19) of AR632719 Sequence 1 from patent U
Query Match      683;  Conservative     0;  Mismatches  606;  Indels  89;  Gaps  7;
Best Local Similarity   49.6%;  Pred. No. 4.2e-28;
Matches  683;
```

```
Db 103267 TATCAATACCAGTAGGATTTTTTCGAGTAGTATCAATCACACGGAATATTAGCCTGTTGGGGATA 103208
          ||||                |||||||||                 ||  ||  ||||| ||  |
Qy   1320 CCAAAAAATTCACAGCTGACAGTCCAGAGCTCCCAAAACCGACGACGGCCATATTTTTT        1377
          ||||     |||||             |||||||     ||||      |||||     ||||||
Db 103207 AGAGATAATCCACACAAGAAAAGACCTGTTTTTGCCAAGCCCGATGATGACTAGCTAATAATATTTTTT 103150
```

Search completed: May 12, 2006, 09:55:03
Job time : 7232 secs

```
Qy    181 GAGTAAAATCTTGCTCAATTATCGGGGGGCATCTTGACCGGATTAAATACACTATCCACCCA  240
Db 104277 CGATGAAAATTTTGCAAGCAGCGCACCATCTCGACCAAAACAAATAACAAATAAGTGTGG     104218

Qy    241 CTGCAATAACATATGCTGACAGCTCACTCAAAATCTTGACCTTTACCAAGCCCACCCAAAAT  300
Db 104217 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCCGTCTCCGCCTAGCAA    104158

Qy    301 CAATGCCAGTGATGATATGCCATAAAACCGTGCCCAAGCCCATCCAATTGCTGCAATGGTTGA 360
Db 104157 CAAATGCAATTTACCCTCAATATA---------AAGCCCAGCCAATGCAGCAACTGTACT   104107

Qy    361 GCCGATGTTTGTACCTTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA  420
Db 104106 CCCCACATTTGTTGCTTTAGAGTCATTAAATCCAACGAATGCCATTAGCTTGATGCACTAA  104047

Qy    421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCGATGGCATCCATGGG  480
Db 104046 TTGAAAACGATGATCTAACCCTTTGAAATGACGAAGTGCCGGTACGAATTGAAATCTAAATT 103987

Qy    481 AAGATTGGGCAGCTGTGCCTAAGGCAGTGCCAGAGATAAGGCATTGAGGTAGGTTAGCTTGCC  540
Db 103986 AATACCTATAGCTTGTTGTGGCCAAATGCTGTTGCTGCCAAAATGGTTCAATATAATTAATTATGGCGACC 103927

Qy    541 TTTGATTTTTTAATTGATGGATTGGTAAAAGTTGTTTTTTTTGATGATATAATGCCATACC  600
Db 103926 AACCAATGTAGCTTCTTCACAAGGTGTAAAATCACTTCATCTTTTACCATTAAATATTGCTT 103867

Qy    601 ATCACCCATCATCAAAATAAAATAAAAAATCTGC-CGTTTGGTGGCTAAGTAAGGCTATTGTTG 659
Db 103866 GCCATTTTCAGTTTTTTAGCCAATATAATCCGCACTACTTTTCCCAAAAGAAACGGGTAATTGTTT 103807

Qy    660 TGATAATGGTTGTTGATTTGGGGGCAACGCCTTGTCAGTGTGGTCAAGCAATTGCTTGGCTAAAT 719
Db 103806 CGCTTGATTTTCGTTTTTCCCCAAAAGTCAGCC------------------------      103775

Qy    720 TGGCATCGTCTTGGTAGATGATGGCGGGTTTGGGCAATTTTCAAATTACGCAATTTTTTGTG  779
Db 103774 ---TATCTTCATTGTTCAACACACTTACTTTAGCATTATGATAAATGGCTAATTTTGCTT   103718

Qy    780 CCAAATAAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGATGAGATTTAAAATTG 839
Db 103717 GGCGATAATCTTCTGAAATCCATAATAGCGATCCCATATGATCTTCACGTCACCGTAANGTTT  103658

Qy    840 TTGCCCCTTGAGCTGCCCAAATTTGAGATATCCTCAAGCTGAAAACTTGACAGCTCAAGCA  899
Db 103657 TCGCTGCCGCGGAGCTTTTAAGCTATAAGTTGTCTCAAGCTGAGAAAGCTCTAGTA      103598

Qy    900 CCACCAAAATCCATATTTTTGGATCGTTAATAATTCAAGTGCGAGGGCCTGCCAATATTACCGC 959
Db 103597 CATAAAGTTCACAAATCTTCATTC---AACAATGACAAAGCGGGAAATCCCAATATTTCCGC   103541

Qy    960 CGACACCAACAAATCATGCCCTGCAATGTTTCGCCTACTAAGTCGTAACAGTGC         1019
Db 103540 CCATACCAACTTTCACACCAGCAGCGTTTCGCCCATTTCATAAACTAAAGTTAGGGTAC     103481

Qy   1020 TTTTGGCATTTGAGCCTGTGATGGGCGGGCCATGAGTGATGTTGTTTGTTGCTTAGATCTC  1079
Db 103480 TTTTACCATTTGAACCTGTAATCCCCACAAGTTGGCTTTTGTCGCCGCGGCA--------    103429

Qy   1080 GTGCTTTGAGTGTATCAAGAACAAATTGAACATCGCCTGATCACAGGAATCGCCTTGAGCTTT 1139
Db 103428 ---------GAATAATTCAATATCGCCGATTACTTCCACTCCCGCTTTAA              103388

Qy   1140 TGGCGCGGACAATACTCGGCGGCTTCTTGGGTTAATGCCGGGGCCTGATGATGATCTGATCGG 1199
Db 103387 GTGCGGGTTTGAATTTCTGGTGTTTTTACCGCAAGCCCTGGGCTAATAACATCATCATATCGC  103328

Qy   1200 CTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAACGCCTGTGGGGCAGTT       1259
Db 103327 TTTCAAGTAACCATTCCTGATTTAAACTACCAGTAATGAAGAGGGATTATTTGTGAGGAAGTT  103268

Qy   1260 TATCTGCCAAGGTAGGATTGGGATTACCCATCGATGATGCTAACTTGTTGTGACCGTGGATTGA 1319
```

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: May 12, 2006, 07:54:39 ; Search time 6041 Seconds
10687.994 Million cell updates/sec
(without alignments)

Title: US-10-672-787-35_COPY_11357_12736
Perfect score: 1380
Sequence: 1 taaagcgtgaaccatgttaa.........acgacggcatattctttga 1380

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 2339541228 residues

Total number of hits satisfying chosen parameters: 82166650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :   EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	%Query Match	Length	DB	ID	Description
1	98.2	7.1	817	10	CZ543630	CZ543630 SRAa-aad5
2	92.8	6.7	709	6	CF306940	CF306940 HDA1-05-
3	92	6.7	892	9	BZ579116	BZ579116 msh2_6205
4	77.2	5.6	555	10	AG943479	AG943479 Drosophil
5	69.8	5.1	1445	9	BZ568959	BZ568959 pacs2-164
6	62.2	4.5	623	10	AG952088	AG952088 Drosophil
7	56.8	4.1	628	10	CL691264	CL691264 PRIO156a_
8	55	4.0	474	8	DR032465	DR032465 bda02027O_
9	52.6	3.8	1171	9	B2559806	B2559806 pacs2-164
10	50.2	3.6	838	10	CZ545153	CZ545153 SRAA-aad6
11	49.8	3.6	572	9	BH795426	BH795426 BMBAC376G
12	49.4	3.6	452	2	BE184674	BE184674 PM4-HT068
13	47.6	3.4	1101	10	AL068607	AL068607 Drosophil
14	47.2	3.4	1101	10	CNS0OL0O	CNS0OL0O Drosophil
15	46	3.3	823	8	DR505933	DR505933 WS02715.B
16	45	3.3	590	1	AU287701	AU287701 AU287701
17	45	3.3	757	8	DR473894	DR473894 WS02960.B
18	45	3.3	907	8	DN922769	DN922769 43863.2 C
19	44.4	3.2	442	8	DR565020	DR565020 WS02632.C
20	44.4	3.2	490	7	CO253110	CO253110 WS00818.B
21	44.4	3.2	794	8	DR473810	DR473810 WS00960.B
22	43.6	3.2	692	9	BH983612	BH983612 ode20g11.
23	43	3.1	1035	10	CNS002D5	CNS002D5 Drosophil
24	42.8	3.1	1101	10	CNS0100X	CNS0100X Drosophil
25	42.4	3.1	1101	10	CNS0039G	CNS0039G Drosophil
26	42.2	3.1	747	10	CZ668937	CZ668937 OM_Ba022
27	42.2	3.1	915	9	AZ543308	AZ543308 ENTGK93TR
28	42	3.0	941	11	CNS05L54	CNS05L54 Tetraodon
29	42	3.0	1022	11	CNS07ANB	CNS07ANB end of
30	41.8	3.0	517	1	AU287702	AU287702 AU287702.
31	41.8	3.0	1101	10	CNS000D1	CNS000D1 Drosophil
32	41.8	3.0	616	9	BH383040	BH383040 AG-ND-127
33	41.6	3.0	714	6	BH374901	BH374901 AG-ND-127
34	41.6	3.0	747	6	CA920035	CA920035 EST637753
35	41.6	3.0	639	8	CX639436	CX639436 UCRPT02.4
36	41.4	3.0	760	9	BH367641	BH367641 AG-ND-127
37	41.2	3.0	378	8	BH382726	BH382726 AG-ND-127
38	41.2	3.0	434	8	T92809	T92809 ye23d06.r1
39	41.2	3.0	563	1	AM020818	AM020818 AM020818
40	41	3.0	726	9	BH047969	BH047969 RPCI-24-3
41	41	3.0	902	9	BQ446218	BQ446218 UI-H-EUI-
42	40.8	3.0	693	10	AZ685437	AZ685437 ENTLE37TF
43	40.8	3.0	939	10	CNS012ZY	CNS012ZY Drosophil
44	40.6	2.9	635	8	DN157692	DN157692 GCN001110
45	40.6	2.9	779	10	CW381115	CW381115 fabb001f0

ALIGNMENTS

RESULT 1
LOCUS   CZ543630   817 bp   DNA   linear   GSS 13-MAY-2005
DEFINITION SRAA-aad51b03.g1 Strongyloides ratti whole genome shotgun library
            (SRAAGSS 004) Strongyloides ratti genomic, genomic survey sequence.
ACCESSION   CZ543630
VERSION   CZ543630.1 GI:64673383
KEYWORDS   GSS.
SOURCE   Strongyloides ratti
  ORGANISM Strongyloides ratti
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Panagrolaimoidea; Strongyloididae; Strongyloides.
REFERENCE   1 (bases 1 to 817)
  AUTHORS   Mitreva,M., McCarter,J.P., Thompson,F., Viney,M., Pape,D.,
            Ritter,E., Martin,J., Wylie,T., Dante,M., Waterston,R.H.,
            Clifton,S.W. and Wilson,R.
  TITLE   Genome Survey sequences from the rat parasitic nematode
            Strongyloides ratti
  JOURNAL   Unpublished (2005)
COMMENT   Contact: Mitreva M
            Washington University in St. Louis
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: nematode@watson.wustl.edu
            Genomic DNA was provided by Fiona Thompson
            (F.Thompson@bristol.ac.uk) and Mark Viney
            (Mark.Viney@bristol.ac.uk) at the University of Bristol, Bristol,
            UK.
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..817
                     /organism="Strongyloides ratti"
                     /mol_type="genomic DNA"
                     /strain="Isofemale line ED321 heterogonic"
                     /db_xref="taxon:34506"
                     /dev_stage="infective larval stage (iL3)"
                     /lab_host="GS10"
                     /clone_lib="Strongyloides ratti whole genome shotgun
                     library (SRAAGSS 004)"
                     /note="vector: pOTW13; Site 1: BstX1; Site 2: BstX1;
                     Strongyloides ratti genomic DNA was randomly sheared,
                     end-repaired and size fractioned to enrich for 2-4 kb
                     fragments. Genomic DNA was provided by Fiona Thompson

```
ORIGIN

Query Match    Best Local Similarity  7.1%;  Score 98.2;  DB 10;  Length 817;
Matches 207;  Conservative  0;  Mismatches 138;  Indels  6;  Gaps  2;

Qy    722 GCATCGTCTTGGTAGATGAGGCGTTTGGCAATTTCAAAAATACGCAATTTTTGTGCC 781
          ||||| ||||||| ||| |||||||||||||||||||| |||||||||||||||||||
Db    395 GCATCTTGTACGTTAACAATGCAGTTTTTAGCATCGTTATAAAATTTTAATTAGCTTCA 336

Qy    782 AAAATAATCACCATGCC---ATCGTGCCGATCAAGATGATCAGGCGGAGAGAGAGATTTAAAATT 838
          ||||||||||||||||    ||||||||||||||||||||||||||||||||||| ||||||||
Db    335 CGATACTGCGCCTAGACCTAATGGGTATCTATCCATATGACTTCACTAATATTTTAAGACT 276

Qy    839 GTTGCCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGC 898
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    275 GTCGCGCGGGTCGCTTCTAAACTTGAAGTGGTTTCTTAATTGGAAGCTGCCAATATTACCG 216

Qy    899 ACCACCAAAATCCATATTTTGGATGCGTTAATAATTCAAGTGCAGGGGTTGCCAATTTACCG 958
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    215 ACGTATAGACG---ATGCCCTGAGTTAGTAATGAAAGCGCTGGGGTACCCAAATATTGCCA 159

Qy    959 CCGACACCAACAATCATGCCTGCAGTGTTTTGCCATCTCGCCCTACTAAATGTCGTAACACGTG 1018
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    158 CCAACACCAACTGAAATACCCGCAGCTTTAGCCAATTCACCAACCAAATGACGTGACGGTA 99

Qy   1019 CTTTTGGCATTGAGCCTGTGATGGCGATGATTGGTGTTGTTTGCTTGA 1069
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db     98 CTTTTACCATTTGAACCTGTAAATGGCGACAAATCCGTGCATCGGCGTTCACGA 48
```

```
RESULT 2
LOCUS       CF306940                      709 bp    mRNA        linear       EST 15-AUG-2003
DEFINITION  HDA1--05-E12.g1 OsHDAC1-overexpressing transgenic rice lambda phage
            cDNA library I (HDA1) Oryza sativa (japonica cultivar-group) cDNA
            clone HDA1--05-E12, mRNA sequence.
ACCESSION   CF306940
VERSION     CF306940.1  GI:33678701
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group)
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 709)
  AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
            Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
  TITLE     Large-scale Sequencing Analysis of Rice ESTs
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Nahm B.H.
            Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
            of Bioscience and Bioinformatics, Myongji University
            Yongin, Kyeonggi, Korea
            Tel: 82 31 330 6193
            Fax: 82 31 321 6355
            Email: bhnahm@bio.myongji.ac.kr.
FEATURES             Location/Qualifiers
     source          1..709
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nackdong"
                     /db_xref="taxon:39947"
                     /clone="HDA1--05-E12"
                     /tissue_type="callus"
                     /dev_stage="proliferated callus on 2N6 media for 2 weeks"
                     /lab_host="E.coli SOLR"
                     /clone_lib="OsHDAC1-overexpressing transgenic rice lambda
                     phage cDNA library I (HDA1)"
ORIGIN
```

```
Query Match    Best Local Similarity  6.7%;  Score 92.8;  DB 6;  Length 709;
Matches 204;  Conservative  0;  Mismatches 142;  Indels  6;  Gaps  2;

Qy    721 GGCATCGTCTTGGTAGATGAGGCGTTTGGCAATTTCAAAAATACGCAATTTTTGTGC 780
          |||||| ||||||| ||| |||||||||||||||||| ||||||||||||||||||||
Db    449 GGCATCATCAGCATTAACCACGCAAACTTCGCGTTTTCCTAAAATGGCAGTTTTTGCTGC 390

Qy    781 CAAATAATCCACCATGCCA---TCGTGCCGATCAAGATGATCAGGCGGAGAGAGATTTAAAAT 837
          ||||||||||||||||||     ||||||||||||||||||||||||||||||||||||||||
Db    389 ACGATACTGTTGTAAACCAAAACGGATAGGCGATCCATAATGGCTTCAGTTCACGTTCAGAAT 330

Qy    838 TGTTGCCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAG 897
          |||||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db    329 GGTCCGCTGCTACCGCCGCTGAGCTGGACGTGGCGGTTTCCAGCTGGAAGGCTCGCAGTTCCAG 270

Qy    898 CACCACCAAAATCCATATTTTGGATGCGTTAATAATTCAAGTGCAGGGGTTGCCAATTTACC 957
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    269 CACGTACAGTTCACACTCATCATCCAGTACCATC---AACGCAGCACGCAGGGGTGTAACAGT 213

Qy    958 GCCGACCAACAATCATGCCTGCAGTGTTTTGCCATCTCGCCCTACTAAATGTCGTAACAGT 1017
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    212 ACCCACCACCAACGTTAACCCCCGCGGCCTTTCGCCAATTCACCCACTAGCGTGGTGGACCGT 153

Qy   1018 GCTTTTGGCATTTGAGCCTGTGATGGCGATGATTGGTGTGTTTGTTTGCTTGA 1069
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db    152 GCTTTTGCCGTTAGAACCGGTAATCGCCACAAATCGGTGCTTGTGTCGCTTCGCGGA 101
```

```
RESULT 3
LOCUS       BZ579116                      892 bp    DNA         linear       GSS 17-DEC-2002
DEFINITION  BZ579116           mbh2_6205.y2 msh Pseudomonas aeruginosa genomic clone mbh2_6205,
            genomic survey sequence.
ACCESSION   BZ579116
VERSION     BZ579116.1  GI:27214177
KEYWORDS    GSS.
SOURCE      Pseudomonas aeruginosa
  ORGANISM  Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1  (bases 1 to 892)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE     Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
  JOURNAL   Bacterial. (2002) In press
COMMENT     Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 206221954
            Fax: 2062169244
            Email: craymond@u.washington.edu
FEATURES             Location/Qualifiers
     source          1..892
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="MSH"
                     /db_xref="taxon:287"
                     /clone="mbh2_6205"
                     /clone_lib="mbh"
                     /note="Environmental isolate. Whole genomic shotgun
                     library."
                     /Class: shotgun.
ORIGIN
```

```
/note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
XhoI; Callus was treated with ABA(20um) for 1hour, cDNA
was inserted into lambda Uni-ZAP XR vector at 5' end with
EcoRI and 3' end with XhoI site. mRNA was derived from
rice Histone Deacetylase overexpression line."

(F.Thompson@bristol.ac.uk) and Mark Viney
(Mark.Viney@bristol.ac.uk) at the University of Bristol,
Bristol, UK. Sequencing by Washington University Genome
Sequencing Center, St. Louis, MO."
```

RESULT 4

```
Query Match     6.7%;  Score 92; DB 9; Length 892;
Best Local Similarity 56.3%; Pred. No. 8e-15;
Matches 192; Conservative   0; Mismatches 146; Indels   3; Gaps   1;
```

```
Qy  721 GGCATCGTCTTGGTAGGATGGATGGCGGTTGGCAATTTCAAAAATACGGCAATTTTGTGC 780
        |||||  ||| | | | |||| ||   |||||||||||| ||||||||||||||||| |||
Db  256 GGCATCGGCGGGATTCACCACGGACCTTGGCAATTTTCAAAATACGGCAATTTTTGGCCAG 315

Qy  781 CAAATAATCCACCATGCCATCGTGTCCGATCAAGATGATCACGGGAGAGATTTAAAATTGT 840
        |||| |  |  |||||  ||| ||   |||| |  |   | |  | |||||||| | ||
Db  316 GTGGTAGTCAGCCATCGCGTCGTAGCGGATCCATATAGGTCTTCGCTGACGTTCAGCACGGT 375

Qy  841 TGCCCTTGAGCTGCCAAAATTTGAGAGATATGCTCAAGCTGTGAAAACTTGACAGCTCAAGCAC 900
        ||||  | | | ||| |  || |  | | |||  |  |||  |||| ||| |||| || |||||
Db  376 CGCCACCTCGGCGCTGAGGCGATCGCAGGTTTCCAGCTGGAAGCTCGACACGTTCAACAC 435

Qy  901 CACCAAAATCCATATTTTGGATCGTTAATATAATTCAAGTGACGGTGCCAATATTACCGCGCC 960
        | |  |  ||| || || |||||      | |  |  ||| ||| ||||||   |||| ||||
Db  436 GTACAGCTCGATGTCTCCGGCC---AGCAGGTCGGACGCGCCGGGGTGCCGAGGTTGCCGCC 492

Qy  961 GACACCAACAAATCAATCATGCCTGCATGTTTTGCCATCTGCCTACTAAGTCGTAACAGTGCT 1020
        ||| || || |  | ||| |||| || | || ||||| | ||| | ||| |||||||||||||
Db  493 GACGGCGACACGCTTGTCCGCGGGCGCCACCGCCATTCGCCCACCAGGGTGGTCACGGGTGCT 552

Qy 1021 TTTGGCATTGTCGTGATGGCGGATGGATTGGTGCGTGTTTTG 1061
        ||| ||| ||| | || || |  ||| || || | | ||||
Db  553 CTTCGGCGTTGGAACCGGTGATGGGGACGATCGGGGNGCCTTCG 593
```

LOCUS   AG943479   555 bp   DNA   linear   GSS 01-FEB-2005
DEFINITION Drosophila ananassae DNA, clone: DNB1-017M23.R.fa, genomic survey
            sequence.
ACCESSION   AG943479
VERSION   AG943479.1 GI:58464900
KEYWORDS   GSS.
SOURCE   Drosophila ananassae
  ORGANISM Drosophila ananassae
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 555)
  AUTHORS   Hattori,M., Toyoda,A., Murakami,K., Kuroki,Y., Fujiyama,A.,
            Toshio,T.K. and Sakaki,Y.
  TITLE   BAC end sequences of library DNB1
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 555)
  AUTHORS   Hattori,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-JAN-2005) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT   BAC end sequences of library DNB1
            Clones are derived from the BAC library DNB1
            For BAC library availability, please contact Masa-Toshi Yamamoto
            (yamamoto@kit.jp).
            This work was done in collaboration with Yamamoto, M-T. Drosophila
            Genetic Resource Center
            Saga Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan
            Tel: 81-75-873-2660 FAX: 81-75-861-0881
            PRIMERS
              Sequencing : R
            LIBRARY
              Vector   : pKS150
              R.Site 1 : SacI
              R.Site 2 : SacI
FEATURES   Location/Qualifiers
     source   1..555
                     /organism="Drosophila ananassae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7217"
                     /clone="DNB1-017M23.R.fa"
                     /clone_lib="DNB1 Drosophila BAC library"
ORIGIN

RESULT 5

```
Query Match     5.6%;  Score 77.2; DB 10; Length 555;
Best Local Similarity 53.3%; Pred. No. 1.4e-10;
Matches 163; Conservative   0; Mismatches 143; Indels   0; Gaps   0;
```

```
Qy  758 TCAAAAATACGGCAATTTTTGTGCCCAAATAATCCACCACATGCCATCGTGTCCGATCAAGAGA 817
        ||||||||| |  |||||||| |  | | |  ||    | |||| |||   || |   ||| ||
Db  426 TCAAAAATATGATGTTTTTGCTCGTGATATATCCTTGCATATATCCTCCATGACGGCTTCTAAATGA 367

Qy  818 TCAGCCGGAGAGAGATTTAAAATGTTTGTGCCCCTTGAGCGTGCCAAAATTTGAGATATGCTCAAGC 877
        || |    | | | |||| |||| ||||  | ||  | |  |   |||| | || | | |||| | |
Db  366 TCTGGGAGTTATATAATTAAGTAAAACATAGAAATATATCAGCCTGTAAAATTCTACGAAACGTTCTAAC 307

Qy  878 TGAAAACTTGACAGCTCAAGCCACGACCACCAAAATCCATATATTTTGGATCGTTAATAATTCAAGT 937
        | || | || || | || |  || || |  |||  | ||| |  || |||| ||| |  || |  |
Db  306 ATATAAGAGGACATCTCTAGAACATATACGCCTTGATCATCCAATAAAAGGAAGAGCACCAGT 247

Qy  938 GCAGGCCGTGCCAATATTACCGCGCGACCACCAAATCATGCCTGCATGTTTTTGCCATCTCG 997
        |||| || | |||| |||| ||||| |||| |  || |||| || || |||||| || | ||
Db  246 GCAGCAGGACCTTAAATTCCCTCCTGCCATCGGCGAGGAAATACCAGCTTCATTTAAAAGAGATGT 187

Qy  998 CCTACTAAGTGTCGTAACAGTGCTTTTGTGCATTTGAGCGTTGTGATGGGCGGATGGATGGTGTG 1057
        ||| | ||| ||||||||| |||||| ||||| |||| | ||||| || |  ||| ||   | |
Db  186 GTTAATAAAACCGTTGTTGTGGCATTTCCCATTTGTGCCTGTGGTCGAACAAACGTGCT 127

Qy 1058 TTTGTT 1063
        |||| |
Db  126 TTTGAT 121
```

LOCUS   BZ568959   1445 bp   DNA   linear   GSS 17-DEC-2002
DEFINITION pacs2-164_8171.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
            pacs2-164_8171, genomic survey sequence.
ACCESSION   BZ568959
VERSION   BZ568959.1 GI:27202799
KEYWORDS   GSS.
SOURCE   Pseudomonas aeruginosa
  ORGANISM Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 1445)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
  TITLE   Whole-Genome-Sequence Variation among multiple isolates of
            Pseudomonas aeruginosa genomic library
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT   Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..1445
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="2-164"
                     /db_xref="taxon:287"
                     /clone="pacs2-164_8171"
                     /clone_lib="pacs2-164"

RESULT 6

```
AG952088
AG952088
```

LOCUS   AG952088   623 bp   DNA   linear   GSS 01-FEB-2005
DEFINITION Drosophila sechellia DNA, clone: DSE1-004A05.F.fa, genomic survey
          sequence.
ACCESSION   AG952088
VERSION   AG952088.1 GI:58473786
KEYWORDS   GSS.
SOURCE   Drosophila sechellia
  ORGANISM Drosophila sechellia
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 623)
  AUTHORS   Hattori,M., Toyoda,A., Murakami,K., Kuroki,Y., Fujiyama,A.,
          Toshio,T.K. and Sakaki,Y.
  TITLE   BAC end sequences of library DNB1
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 623)
  AUTHORS   Hattori,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-JAN-2005) Masahira Hattori, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
          (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.jp/,
          Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT   Clones are derived from the BAC library DSE1
          For BAC library availability, please contact Masa-Toshi Yamamoto
          (yamamoto@xit.jp).
FEATURES   Location/Qualifiers
     source   1..623
                  /organism="Drosophila sechellia"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:7238"
                  /clone="DSE1-004A05.F.fa"
                  /clone_lib="DSE1 Drosophila sechellia BAC library"

ORIGIN

Query Match   5.1%; Score 69.8; DB 9; Length 1445;
Best Local Similarity 55.0%; Pred. No. 2.5e-08;
Matches 137; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```
Qy    327 CCGTGCCAAGCCCATCAATTGCTCTGCAAATGGTTGAGCCGGATGGTTTGTACCTTTTGAATCAT 386
              ||||||||||| |||||||| |||||||| ||||| |||| |||  | ||| || |||||
Db     87 CGGCACCCACGCCCTCGATCGCCCCGCCAGGGCGCGGGCCCGAGGTTGGTTGGCCTTGGAATCGT 146

Qy    387 TAAAATATGCCTTGCTATCAAATACATCAATATATTCACAGCGGATGCGGGTAGCCCTTTTGA 446
              ||| | ||| || ||| |||| |||| || || ||  ||| || ||| ||| || |
Db    147 CGTAGTAGCTCACCGCCCTCGCCGTACCCACTGCGGCTGGCGGATGAGCCAGGCGCCGGAAA 206

Qy    447 AAGTCTTAAGGGTATCAAGCATGGCATTCCATGGGAAGATTGGGCACGTGTCGCCTAAGGCAA 506
              |||||| || |||| || |||| | ||  |  ||| || ||||  ||| | |||| |||
Db    207 ACGCCTTCAGGCGCCCGCGAGCGATGGGCCCGGACGCACCGCATGGCCCACGCGCCA 266

Qy    507 GTGGAGATAAGGCAATTGAGTGGGTTATGCTTGCTTTGATTTTTAATTGATGGATTGGGTA 566
              ||| ||| |||||| || |||| | |||  ||| |||| || |||| |||| | |||
Db    267 GCGGCGCGACGGCGTTGGAATAGTTGTGGGCGCCACAGGACTTCAGTTCGCCACCGGCA 326

Qy    567 AAAGTTTGT 575
              ||| ||||
Db    327 GCAGCTTGT 335
```

ORIGIN

RESULT 7

```
CL691264/c
CL691264
```

LOCUS   CL691264   628 bp   DNA   linear   GSS 10-JUL-2004
DEFINITION PRIO156a_A05_2 - PRIO156a.BR (628) Mixed stage fosmid library of P.
          pacificus var. California Pristionchus pacificus genomic, genomic
          survey sequence.
ACCESSION   CL691264
VERSION   CL691264.1 GI:50213172
KEYWORDS   GSS.
SOURCE   Pristionchus pacificus
  ORGANISM Pristionchus pacificus
          Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
          Neodiplogasteridae; Pristionchus.
REFERENCE   1 (bases 1 to 628)
  AUTHORS   Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  TITLE   AppaDB: an AcedB database for the nematode satellite organism
          Pristionchus pacificus
  JOURNAL   Nucleic Acids Res. 32 (1), D421-D422 (2004)
  PUBMED   14681447
COMMENT   Contact: Sommer RJ
          Evolutionary Biology
          Max-Planck-Institute for Developmental Biology
          Spemannstr. 37-39, Tuebingen D-72076, Germany
          Tel: 0049707160137
          Fax: 00497071601498
          Email: ralf.sommer@tuebingen.mpg.de
          This library was generated at Caltech, Pasadena, USA and end
          sequenced at Vancouver, Canada.
          Seq primer: T7
          Class: fosmid ends.
FEATURES   Location/Qualifiers
     source   1..628
                  /organism="Pristionchus pacificus"
                  /mol_type="genomic DNA"
                  /strain="California"

ORIGIN

Query Match   4.5%; Score 62.2; DB 10; Length 623;
Best Local Similarity 51.0%; Pred. No. 3.1e-06;
Matches 176; Conservative 0; Mismatches 163; Indels 6; Gaps 1;

```
Qy    703 GCAATGCTTGCTAAATTGCCATCGTCTTGGTAGAATGAATGGGGGGTTTGGCAATTTTCAAA 762
              |||  ||| ||||||| |||||  ||||| || ||| ||| || |||  ||| || ||||
Db    285 GAAATATATCAGCCAGTAATCTCCATTGTCTCACATCCTATTACGGCAACCTCACTACGTTTTAT 344

Qy    763 AATACGCCAATTTTGTGCCAAATAATCCACCACGTCGTCGCCGGATCAAGAGATGATCAGC 822
              ||||| | ||||| |||  | ||| |||  ||| |||  | |||  || ||| ||  |||
Db    345 CAGTTTAATTTAGTTGGCTARTGTAGTTCTCCATACTTCACATCCARTGTAGACTGG 404

Qy    823 GGAGAGATTTAAAATTGTTGCCCCTTGAGCTGTGCCAAAATTGAGATATGCTCAAGCTGAAA 882
              ||||| ||||||| || || |||| ||| |||| ||||| || ||| ||| |||||||||
Db    405 TGTAATATTGAGCAGTGCTGAAAGTCCACGTTAATTCATTCATCACACTCCAATTGAAA 464

Qy    883 ACTTGACAGCTCAAGCACCACACCAAATCCATATTTTGGATGTTAATAATTCAAGTGCAGG 942
              ||| | |||||||| ||| ||| |||||| ||||| ||| ||| || || |||||||||
Db    465 AGAGGAGAGTTCGATTACAATAAATTTCTGCATCTC------TTCTAGAATCCAAAAACAGG 518

Qy    943 CGTGCCAATATTACCGCCGGACCAACAAATCATGCCCTGCGATAGTTTTGCCAATCTCGCCTAC 1002
              ||| || ||||| |||| || || ||| |||| | ||| ||  ||||| |||| || || ||
Db    519 AACGCCTAAATTTCCACCAATAGCTACTTTTTCCCTGCAGAATTTTAATATGTGCCTAT 578

Qy   1003 TAAATGTCGTAACAGTGCCTTTTGGCATTGAACCTGTGATGGCGAT 1047
              |||| ||| ||| || ||| ||||  ||| || |||  ||| |||
Db    579 TAGTGATGTAGTGGTTGGTTGATTACCAATTTGTTCCTGTGACGCCTAT 623
```

PRIMERS

  Sequencing : F

LIBRARY

  Vector   : pKS150
  R.Site 1   : SacI
  R.Site 2   : SacI

ORIGIN
/db_xref="taxon:54126".
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pEpifos-5 Fosmid vector"

Query Match   48.0%; Score 56.8; DB 10; Length 628;
Best Local Similarity   4.1%; Pred. No. 0.00011;
Matches 246; Conservative 0; Mismatches 242; Indels 24; Gaps 2;

Qy   1 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGGCGGTGCTTTGACTGGGCTGCCAA 60
           ||| ||| | || | || ||| || ||| || | || ||| ||| ||| ||| |||
Db   494 TAACTCCTTCGCCAGACGGGCGCAAATCATTCGCCTCGTTGTTCAAAGTTCTTGAACTGATC 435

Qy   61 CGGGGCGTAGGCGCCGCGCCAGCTGCGCCGGTCACGACCGGAAACAAATACAGTACGTTATC 120
           || | | || || |||| || |||| || | || |||||| ||| |||||||| |
Db   434 AAGGCTGGGCACAGGCTGGGGAGGAGGAGAACCAT------------AICGCCCGGCTG 390

Qy   121 TACTGCCCGGATCAAGTGTACCAACCTGATATATTTTTAGATGCCAAAGTGCATCACCTTT 180
           || ||||| || || |||| |||| || || |||| || ||||||||| ||| | || |
Db   389 AACACGCGGGAGCCAGGAGCCCATCCCGTCCATAGTTTCGGTTGTTGTTCTGCCACTTC 330

Qy   181 GAGTAAATCTTGCTCAAATTATCGGGGCATCTTGACCGGATTAAAATACACTATATCCACCCA 240
           |||| | | || || || |||| |||| ||||| || || ||| |||| || |||||||
Db   329 CGGGGGTAGGCGCCGCGCCAGCTGCGCCGGTCACGACCGGAAACAAATACAGTACGTTATC 270

Qy   241 CTGCATACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCACCAAAAT 300
           || ||| | | ||| || |||| || || | ||||| || |||| |||||||||||||
Db   269 GCCATTCAGTAACCGCCAGTGGCTAAAGTCCCGCCGATTTACCATCGGCTCCCAGCAA 210

Qy   301 CAATCCCAGTGATGATATGCCATAAACCGTTGCCAAGCCCCATCAAATTGCTCAATGGTTGA 360
           ||| ||||| || ||| || || | |||||| |||| ||||| | ||| || | |||
Db   209 CAAATGCAGTGTGCCGT-------CTACGTGCAGGCCATTCAGCGCCGCCTTCCGTACT 159

Qy   361 GCCGGATGTTTGTACCTTTTGAATCATTAAAATAATGCCTTGCTATCAAATATCATCAATATA 420
           ||||| || || |||| || ||||| || || ||||| || || ||| || ||||||||
Db   158 GCCGACGTTGGTCGCGTTTCGAAATCGTTAAATCCAACGTACCGCCGGTTATGCTCCAGCACAC 99

Qy   421 TTCACAGGGAATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGGCATCCCATGGGG 480
           ||||| |||| || || | ||||| | |||||| | |||||| |||| | || ||||
Db   98 TTCAAAGGCCATGGCGCAGACCAGTGAATGTGGTTAACGCTTTCAGGCTGCTGGGCACGCGG 39

Qy   481 AAGATTGGGAGCTGTGCCTAAGGCAAAGTGCAG 512
           ||||
Db   38 TAACCCTGCAGCATCTGCCAGCGCCAGCGCCG 7

RESULT 8
DR032465/c
LOCUS   DR032465   474 bp   mRNA   linear   EST 31-MAY-2005
DEFINITION bda02070D15.ab1_061 Antrodia cinnamomea cDNA, mRNA sequence.
            cinnamomea cDNA, mRNA sequence.
ACCESSION   DR032465
VERSION   DR032465.1 GI:66834715
KEYWORDS   EST.
SOURCE   Antrodia cinnamomea
ORGANISM   Antrodia cinnamomea
            Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
            Aphyllophorales; Antrodia.
REFERENCE   1 (bases 1 to 474)
AUTHORS   Tzean,S.S., Hwang,S.T., Song,H.C., Liou,R.F., Yehk,H.H. and
            Shiao,M.S.
TITLE   Analysis of expressed sequence tags from Antrodia cinnamomea, a
            polypora causing brown rot of Cinnamomum kanehirai
JOURNAL   Unpublished (2005)
COMMENT   Contact: Tzean SS
            Department of Plant Pathology and Microbiology
            National Taiwan University
            NO.1, Section 4, Roosevelt Road, Taipei 10617, Taiwan
            Tel: 886 2 3366 4595
            Fax: 886 2 2362 0639
            Email: sst@ntu.edu.tw
            PCR Primers

FORWARD: 5' GTAAAACGACGGCCAGT 3'.
BACKWARD: 5' CACACAGGAAACAGCTATGACCAT 3'.
FEATURES   Location/Qualifiers
     source   1..474
                     /organism="Antrodia cinnamomea"
                     /mol_type="mRNA"
                     /strain="SMV 1"
                     /db_xref="taxon:279009"
                     /sex="Heterokaryon"
                     /cell_type="Mycelium"
                     /lab_host="Cinnamomum kanehirai"
                     /clone_lib="Antrodia cinnamomea cDNA Library"
                     /note="Vector: ZAP Express Vector"
ORIGIN

Query Match   54.1%; Score 55; DB 8; Length 474;
Best Local Similarity   4.0%; Pred. No. 0.00035;
Matches 112; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

Qy 1171 AATGCCGGGGGCTGATGATGATCTGATCGGCTGCTGAGCTTAAGAGTTCACTGTCCAAGCTGCC 1230
           || | | ||||| || | || ||||| | |||| || || ||| ||| || ||| |||||
Db   453 AATACCGGGACTGGCGACAATCAGATCTCGCGCCGCCATCAGCCATTCAGCTGCTGGGCACGCGG 394

Qy 1231 AAAATGACAATGAAACGCCTGTGGGCAGCAGTTTTATCTGCCAAGGTAGGATTGGGATTACCATC 1290
           ||||| ||| | | ||| || || | || ||| | || |||| ||| | | ||||| ||
Db   393 CGTGTGGCGTTCTACGGCTTCGGGTTAATTTATTCCAGGCCGGTGTGTCATACGCGGTATC 334

Qy 1291 GATGATGCTAACTTTGTGACCGTGATTGACCAAAAAAATTCACAGCTGACAGTCCAGAGCT 1350
           |||| || ||| | ||| || ||||| | |||| ||||| | | | || ||||||
Db   333 CATAACGCGCCGCGGTCACACCGCGACGAGGAAAAGTCCACGCACGACGAAAGCCCGGTGAG 274

Qy 1351 GCCCAAACCGACGACGGCATATTTTTT 1377
           |||||| || | | |||||||||||
Db   273 GCCCAGCCGGATAAATGACGACATTTTT 247

RESULT 9
BZ559806/c
LOCUS   BZ559806   1171 bp   DNA   linear   GSS 17-DEC-2002
DEFINITION pacs2-164_1863.x1 pacs2-164 Pseudomonas aeruginosa genomic clone
            pacs2-164_1863, genomic survey sequence.
ACCESSION   BZ559806
VERSION   BZ559806.1 GI:27176698
KEYWORDS   GSS.
SOURCE   Pseudomonas aeruginosa
ORGANISM   Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 1171)
AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE   Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa library
JOURNAL   J. Bacteriol. (2002) In press
COMMENT   Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2068857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..1171
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="2-164"
                     /db_xref="taxon:287"
                     /clone="pacs2-164_1863"
                     /clone_lib="pacs2-164"
                     /note="clinical isolate 2-164 Whole genomic shotgun
                     library."
ORIGIN

RESULT 10

```
Query Match    Best Local Similarity  49.0%;    Score 52.6;  DB 9;  Length 1171;
Matches 170;  Conservative  0;  Mismatches 174;  Indels  3;  Gaps  1;  Pred. No. 0.0022;

Qy   177  CTTTGAGTAAATCTTTGCTCAATTATCGGGCGCATCTGGTTGACCGGATTAAATACACACTATCCA  236
Db   396  CGTTGCCCAGTGCCTGGTCAATCAGCTCAGTGTCACGCACAAGAGCAGTACCACCGCCCGGC      337

Qy   237  CCCACTGCATAAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAAGCCCACCCA     296
Db   336  AGAAGCGCGCGGCGGCTCCGCCAGGTCATGAAAAATCGGGCGGCCCTTGCCGTCTCCGGCCG     277

Qy   297  AAAATCAAATGCCAGTGATGATATGCCATAAACCGGTCCCAAGCCCATCAACAAATTGCTTGCAATGG  356
Db   276  CGATCAGCACCAG---CTTGCCGTCGATGTCGGCAACCAGCACCCCCTTCGATCGGCCCGCCCAGGG   220

Qy   357  TTGAGCCCGATNGTTTTGTTACCTTTTGAAATCATTAAAATATGCCTTGCCTATCAAATATCATCAA   416
Db   219  CGGCGCCGACGTTGGTGGCCTTTGAAATCGTTGTAGCTAGCTCACCGCCCTCTGCCGGTA         160

Qy   417  TATATTCACAGCGATGCGGTAGCCCCTTTGAAAGTCTTAAGGGGTATCAAGGCCATGGCCATCCA    476
Db   159  CCCACTGCGAGCAAATGAGCCAGTCCGCAAGAGCGCTTTCAGCCGGCCGGAGCGCATGGCCGTAAA   100

Qy   477  TGGGAAGAATTGGCAGCTGTCCCTAAGGCAAAGTGCAAGATAAAGGCAATTG                 523
Db    99  ACGGCAGGCCGGACCGCCATGGCCCAGTCGCACGCTCGGCGAGCCGCGTTG                   53
```

```
LOCUS       CZ545153    838 bp    DNA    linear   GSS 13-MAY-2005
DEFINITION  SRAA-aad60c01.g1 Strongyloides ratti genomic whole genome shotgun
            sequence. genomic survey sequence.
ACCESSION   CZ545153
VERSION     CZ545153.1  GI:64677289
KEYWORDS    GSS.
SOURCE      Strongyloides ratti
  ORGANISM  Strongyloides ratti
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Panagrolaimoidea; Strongyloididae; Strongyloides.
REFERENCE   1  (bases 1 to 838)
  AUTHORS   Mitreva,M., McCarter,J.P., Thompson,F., Viney,M., Pape,D.,
            Ritter,E., Martin,J., Wylie,T., Dante,M., Waterston,R.H.,
            Clifton,S.W. and Wilson,R.
  TITLE     Genome Survey sequences from the rat parasitic nematode
            Strongyloides ratti
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Mitreva M
            Washington University in St. Louis
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: nematode@watson.wustl.edu
            Genomic DNA was provided by Fiona Thompson
            (F.Thompson@bristol.ac.uk) and Mark Viney
            (Mark.Viney@bristol.ac.uk) at the University of Bristol, Bristol,
            UK.
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1..838
                     /organism="Strongyloides ratti"
                     /mol_type="genomic DNA"
                     /strain="Isofemale line ED321 heterogonic"
                     /db_xref="taxon:34506"
                     /dev_stage="infective larval stage (iL3)"
                     /lab_host="GS10"
                     /clone_lib="Strongyloides ratti whole genome shotgun
                     library (SRAGSS 004)"
                     /note="Vector: pOTW13; Site 1: BstXI; Site 2: BstXI;
                     Strongyloides ratti genomic DNA was randomly sheared,
```

```
ORIGIN
            end-repaired and size fractioned to enrich for 2-4 kb
            fragments. Genomic DNA was provided by Fiona Thompson
            (F.Thompson@bristol.ac.uk) and Mark Viney
            (Mark.Viney@bristol.ac.uk) at the University of Bristol,
            Bristol, UK. Sequencing by Washington University Genome
            Sequencing Center, St. Louis, MO."
```

RESULT 11

```
Query Match    Best Local Similarity  62.2%;    Score 50.2;  DB 10;  Length 838;
Matches 79;  Conservative  0;  Mismatches 48;  Indels  0;  Gaps  0;  Pred. No. 0.01;

Qy   946  GCCAATATTACCGCGCGACACCAACAATCATGCCGTCGCATGTTTTGCCATCTCGCCTACTAA  1005
Db    76  GCCAATAATTTCCGCGCCTAAAACCAACARTCAACCTTCCATTTTTCAGGATGTGAATAGATTAA  135

Qy  1006  TGTCGTAACAGTGCTTTTTGGCATTTGAGCGTCTGTGATGGCCGATGATTGGGTGTGTTTGTTGC  1065
Db   136  AGACGTTGTCGTTGTTTTTCCGTTGCTTCCTTCGTGTGATCGGCAAATGANTTTCGCATCCGTAAA  195

Qy  1066  TTGAGTT  1072
Db   196  TTCAGAT   202
```

```
LOCUS       BH795426    572 bp    DNA    linear   GSS 09-APR-2002
DEFINITION  BMBAC3376G09?7 PSU Brugia malayi Genomic Bac Library 3 Brugia malayi
            genomic, genomic survey sequence.
ACCESSION   BH795426
VERSION     BH795426.1  GI:20043758
KEYWORDS    GSS.
SOURCE      Brugia malayi
  ORGANISM  Brugia malayi
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Brugia.
REFERENCE   1  (bases 1 to 572)
  AUTHORS   Whitton,C., Daub,J., Quail,M., Hall,N., Foster,J., Ware,J.,
            Ganatra,M., Slatko,B., Barrell,B. and Blaxter,M.
  TITLE     A genome sequence survey of the filarial nematode Brugia malayi:
            repeats, gene discovery, and comparative genomics
  JOURNAL   Mol. Biochem. Parasitol. 137 (2), 215-227 (2004)
   PUBMED   15383292
COMMENT     Contact: Blaxter ML
            Institute of Cell, Animal and Population Biology
            University of Edinburgh
            Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
            3JT, UK
            Tel: +44 131 650 6760
            Fax: +44 131 670 5450
            Email: mark.blaxter@ed.ac.uk
            Sequenced from the Brugia malayi BAC library constructed by Claire
            Whitton and Dr Mike Quail. The sequence was generated by The
            Pathogen Sequencing Unit, The Sanger Institute, Cambridge, UK in
            collaboration with Mark Blaxter, ICAPB, University of Edinburgh,
            Edinburgh, UK
            Seq primer: T7 (TAATACGACTCACTATAGGG)
            Class: BAC ends.
FEATURES             Location/Qualifiers
     source          1..572
                     /organism="Brugia malayi"
                     /mol_type="genomic DNA"
                     /strain="TRS"
                     /db_xref="taxon:6279"
                     /sex="Mixed (male and female)"
                     /tissue_type="whole parasite"
                     /dev_stage="microfilaria (L1)"
                     /clone_lib="Brugia malayi Genomic Bac Library 3"
                     /note="Vector: pBACe3.6; Site_1: BamH I; Brugia malayi
                     genomic DNA was partially cleaved with Sau3A I and size
                     fractionated. 7,392 clones were generated with mean insert
                     size ~48 kbp. The library was constructed by Claire
```

ORIGIN
           Whitton, Blaxter Nematode Genetics Lab, University of
           Edinburgh, UK, and Dr Mike Quail, The Sanger Sequencing
           Unit, The Sanger Centre, Cambridge, UK."

Query Match   3.6%; Score 49.8; DB 9; Length 572;
Best Local Similarity 58.4%; Pred. No. 0.012;
Matches 87; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy   924 TTAATAATTCAAGTGCAGGGCGTGCCAAATATTACCGCGCGACACCAACAATCATGCCTGCAT 983
Db   206 TTTCTAGATCCAAAATAGGAACACACTTAAATTTCCACCAATAGCTACTTTTTTTCCCCGCAG 265

Qy   984 GTTTTGCCATCTCGCCTACTTAATATGTCGTTAACAGTGCTTTTTGGCATTTGAGCCTGTGGATGG 1043
Db   266 ATTTTAATATGTGACCTATTAATGACGTAGTGGTTGATTCCCATTCCGTTCCTGTGACGC 325

Qy 1044 CGATGAATTGGTGTGTTTGTTTGTTGTTGCTTGAGTT 1072
Db   326 CTACAATTTTCTGGTTCTTAGCTTTAGCT 354

RESULT 12
LOCUS   BE184674   452 bp   mRNA   linear   EST 22-JUN-2000
DEFINITION PM4-HT0688-050500-002-d10 HT0688 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BE184674
VERSION   BE184674.1 GI:8663858
KEYWORDS   EST.
SOURCE   Homo sapiens (human)
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 452)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE   Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
  COMMENT   Contact: Simpson A.J.G.
            Ludwig Institute for Cancer Research
            Laboratory of Cancer Genetics
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=PM4-HT0688-050
            500-002-d10&t3=2000-05-05&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 26
            High quality sequence stop: 89.
FEATURES   Location/Qualifiers
   source   1..452
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone_lib="HT0688"
            /dev_stage="Adult"
            /note="Organ: head neck; Vector: puc18; Site_1: SmaI;
            Site_2: SmaI; A mini-library was made by cloning products
            derived from ORESTES PCR (U.S. Letters Patent application
            No. 196,716 - Ludwig Institute for Cancer Research)
            profiles into the pUC 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
ORIGIN

ORIGIN
            low stringency conditions."

Query Match   3.6%; Score 49.4; DB 2; Length 452;
Best Local Similarity 50.5%; Pred. No. 0.014;
Matches 145; Conservative 0; Mismatches 141; Indels 1; Gaps 1;

Qy   332 CCAAGCCCATCAATTGCTGCAATGGAGGTTGAGCCGGATGGTTGTACCTTTTGAATCATT-AAA 390
Db   435 CCAGGCCTGCGGATGGCCTCAGGGCCGGCACCGAGCGTTGGTGGCTTTGGAGTCATTGTTA 376

Qy   391 ATATGCCTTGCTATCAATATCATCAATATATCAATATATCACACGACGTGCGGATGGTAGCCCTTTGAAAGT 450
Db   375 ATAGCTGACTTCATTGAGATCCGGCACCATTGGCACGGATGCTCAAGCCCCGGCATAGT 316

Qy   451 CTTAAGGGTATCCAAGCACATGGCATCCATTGGGAAGAAGAGATTGGGCACGCTGTGCCTTAAGGCCAAGTGC 510
Db   315 GCGCAGGCTCGCGAGCATGGGCGTCGAATAGGCAGCCCAAACCCGCTGTGACCGAGTGCCAAGGC 256

Qy   511 AGATAAGGCATTGAGTAGGTTATGCTTGCCCTTTGATTTTTAAATTGATGGATGGGTAAAAG 570
Db   255 CGCCAGGGCAATTGGACTGGTTGTTGTGCGCCACGGATTGTTCGGTACCGGCATCGG 196

RESULT 13
LOCUS   CNS00LO0   1101 bp   DNA   linear   GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC:
            BACR32D23 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL068607
VERSION   AL068607.1 GI:4958689
KEYWORDS   GSS.
SOURCE   Drosophila melanogaster (fruit fly)
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE   Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
  COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES   Location/Qualifiers
   source   1..1101
            /organism="Drosophila melanogaster"
            /mol_type="genomic DNA"
            /db_xref="taxon:7227"
            /clone="BACR32D23"
            /clone_lib="RPCI-98"
            /note="end : TET3"
ORIGIN

RESULT 14

LOCUS   CNS002MK/c

Query Match   3.4%; Score 47.6; DB 10; Length 1101;
Best Local Similarity   18.8%; Pred. No. 0.061;
Matches   89; Conservative 178; Mismatches 205; Indels   1; Gaps   1;

```
Qy   398 TTGCTATCAATATCATCAATATATATTCACAGCGATCGCGGTAGCCCTTTGAAAGTCTTAAGG 457
             ||::| |::::      | ||  ::|::: : |||    | | : || ||||  ||||||
Db  1081 KTTKGTTKTDDDKNAKVDAAKARTDDDKKTKAKTGWRGKKKKTTTTTKTTKTTKTKTDWAAAAAW 1022
```

```
Qy   458 GTATCAAGACATGGCATCCATGGGAGAGATTGGCAGCTGTGCCTAAGCCAAGTGCAGATAAG 517
            ::|:: :|||   |   ::||  :  || ||::  | :| |:: ||:: | ||
Db  1021 AKTKHKTDTTKGAKTKTTKKGKTNWTKTKKCCKKTSKRKGDAKKAAAAADDKKTDKARAKK 962
```

```
Qy   518 GCATTGAGTAGGTTATGCTTGCCTTTGATTTTAATTGATGGATTGGTAAAGTTTGTTT 577
           | ||| : :||  |  | ||  ||  ||  :  ||  ||:: | ||||:|:||| |||
Db   961 KGKKKKGRKWKKGGKGKKKTKDGTKKKKTTGKWDDDAWMDWWADTDWKKAGGGKGKKDA 902
```

```
Qy   578 TTTTGATGATATAATGCCAATACCATCACCATCAAAATAAAAATCTGCCGTTTCGG 637
           ::|||:| |||| ||  ||||  :||||||::| | |||| : ||  || ||||
Db   901 WTDTWDWKDAWKKGDKKDTKKKKKKGGKAKGRAAAAAADRKGAAAADGDAAAAKTKKTKKAK 842
```

```
Qy   638 TGGCTAAGTAAGCTATATTGTTGTTGTGATAAATGGTGTGATTTTGGGGCCAAGCGTTGTCAGTGTGTG 697
           |   |||| :| ||: ::||:| ||| | ::| || | ||||  ||| :||: |::|: ||
Db   841 TWKGHCKKGGGWGRKGTGRGTKGGKGGTKAGDTKKKKGKRKKKKG 782
```

```
Qy   698 GTCAAGCAATTGCTTGGCTAAATTGGCATCGTCTTCGGTAGATGATGGCGGTTTGGCAATTT 757
             :||| |||| |:||| ||||| |||  |: ::||| | ||| || ||:  ||| ||:||
Db   781 KTGTAWKTKTTGKKKKKGKKGKKDWTKGTKWDTKTKTKIKGKKTKSGSGGGKKKTKYKGTKK 722
```

```
Qy   758 TCAAAAATACGCAATTTTTTTGTGCCAAATAATCCACCATGCGATGCCATCGTGCCGATCAAGATGA 817
           :||| || || | |||  :| |:|| |  |||| ||::|| |||  :|  ||| | || |  | |
Db   721 T-WAAAWKAKKTKKTTKKGTAKTADKRTKKKKDKDKGKTTGKWKKKKTGKAAAWGTGDTR 663
```

```
Qy   818 TCAGCGCGGAGAGATTTAAAATTGTTTGCCCCTTGAGCTGCCCAAATTTGAGATATG 870
            || || ||:||| |||||||||| |::|   :| |  | | |  |  ||||:|:|
Db   662 TKDKKKKAWAAWNDTTDGKTKAKKADTKADTGKKGAKAATKAAKAKGWDADADG 610
```

```
Qy   831 TTAAAATTGTTTGCCCCTTGAGCTGCC 856
           || :  |||||||:||||||||:||:|
Db   705 TRRRGCGSCCGGCGCCCCTTTTCKGTHCY 680
```

ORIGIN   /note="end : SP6"

RESULT 15

LOCUS   DR505933/c

Query Match   3.4%; Score 47.2; DB 10; Length 1101;
Best Local Similarity   36.9%; Pred. No. 0.079;
Matches   76; Conservative 42; Mismatches 88; Indels   0; Gaps   0;

```
Qy   651 TATTTGTTGTGATAATGGTGTGATTTTGGGGGCAACCGCTTGTCAGTGTGGTGGGTCAAGCAATTGCT 710
           :|| |||| |: ::: | |:| ||||||||| |||| || ||  ||||| |||  || | || :|||
Db   885 TTTTTTWAGAAAAAAGCCGYGGRRGKGGGGGGGCAACGSCTKGGTTGCGGTTGGGSGRSMAATTTTK 826
```

```
Qy   711 TGGGCTAAATTGGCATCGTCTTCGGTAGATGATGGCGGCGGGTTTGGGCAATTTCAAAAATACGCA 770
           | ||  || ||||||  ||  |::| |:||  |::||   ||: ||| || |::  |:|| | ||
Db   825 KTTTTAKKDGGGAAGGGATYKGGWAWTTTTAKKGGTWKKKKTTTGTTDDTWAAAAKTAMCSG 766
```

```
Qy   771 ATTTTTGTGCCAAATAAATCCACCATGGCCATGCCATCAAAGATGATCAGCGGAGGAGGAGAT 830
           |||||||::|| |  || |||  ||| |||| |:||| |||| :|| || || ||||:||||
Db   765 WTTTTTYTYCYTTTYTWWCACRAAGSCCTTCWTCCSKTYWWTKWTGKTCRGVCATGTKRG 706
```

```
Qy   831 TTAAAATTGTTTGCCCCTTGAGCTGCC 856
           || :  |||||||:||||||||:||:|
Db   705 TRRRGCGSCCGGCGCCCCTTTTCKGTHCY 680
```

LOCUS   DR505933   823 bp   mRNA   linear   EST 07-JUL-2005

DEFINITION WS02715.BR.P20 SS-IL-A-FL-14 Picea sitchensis cDNA clone
            WS02715_P20 3', mRNA sequence.

ACCESSION   DR505933

VERSION   DR505933.1 GI:70249953

KEYWORDS   EST.

SOURCE   Picea sitchensis (Sitka spruce)

ORGANISM   Picea sitchensis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferophyta; Coniferopsida; Coniferales; Pinaceae; Picea.

REFERENCE   1 (bases 1 to 823)

AUTHORS   Ralph,S., Kolosova,N., Oddy,C., Cooper,D., Butterfield,Y.,
            Kirkpatrick,R., Liu,J., Palmquist,D., Stott,J., Barber,S., Yang,G.,
            Babakaiff,R., Brown-John,M., Chand,S., Featherstone,R., Mason,P.,
            Mayo,M., Moran,J., Olson,T., Wong,D., Friedmann,M.F., Ritland,C.E.,
            Siddiqui,A., Holt,R., Jones,S., Marra,M., Ellis,B.E., Douglas,C.,
            Ritland,K. and Bohlmann,J.

TITLE   The spruce transcriptome: Analysis of expressed sequence tags from
            multiple cDNA libraries

JOURNAL   Unpublished (2005)

COMMENT   Contact: Joerg Bohlmann
            Genome BC forest genomics program
            University of British Columbia
            Michael Smith Laboratories, 6174 University Boulevard, Rm. 237,
            Vancouver, British Columbia, Canada, V6T 1Z3
            Tel: 1-604-822-0282
            Fax: 1-604-822-2114
            Email: bohlmann@msl.ubc.ca
            Plate: WS02715 row: P column: 20
            High quality sequence stop: 831
            POLYA=Yes.

FEATURES   Location/Qualifiers
     source   1..823
                     /organism="Picea sitchensis"
                     /mol_type="mRNA"
                     /cultivar="FB3-425"
                     /db_xref="taxon:3332"
                     /clone="WS02715_P20"
                     /sex="Hermaphrodite"
                     /lab_host="E. coli DH10B cells"
                     /clone_lib="SS-IL-A-FL-14"
                     /note="Organ: Green portion of the leader tissue from one
                     year old clonal trees grown under greenhouse conditions in
                     standard potting soil mixture; Vector: pBluescript II SK
                     (+) XR; Site 1: XhoI (5' end of cDNA); Site 2: BamHI (3'
                     end of cDNA).; 8. mRNA was isolated from each tissue
                     source independently and equal quantities of mRNA from

---

LOCUS   CNS002MK   1101 bp   DNA   linear   GSS 26-JUL-1999

DEFINITION Drosophila melanogaster genome survey sequence SP6 end of BAC
            BACN02P17 of DrosBAC library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.

ACCESSION   AL097862

VERSION   AL097862.1 GI:5609473

KEYWORDS   GSS.

SOURCE   Drosophila melanogaster (fruit fly)

ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.

REFERENCE   1 (bases 1 to 1101)

AUTHORS   Genoscope.

TITLE   Direct Submission

JOURNAL   Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr)
            - Web : www.genoscope.cns.fr)

COMMENT   Determination of this BAC-end sequence was carried out as part of a
            collaboration with the European Drosophila Genome Project (EDGP) -
            http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
            library (Dros BAC) was made by Alain Billaud at CEPH (Centre
            d'Etude du Polymorphisme Humain) with funding provided by a MRC
            project grant. The DNA was prepared from embryos by Alain Bucheton
            and Genevieve Fayan. It has been constructed in the vector
            pBeloBAC11.

FEATURES   Location/Qualifiers
     source   1..1101
                     /organism="Drosophila melanogaster"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7227"
                     /clone="BACN02P17"
                     /clone_lib="DrosBAC"
                     /plasmid="pBeloBAC11"

ORIGIN

each tissue were then pooled. cDNA was prepared from 20 micrograms of mRNA according to the full-length cDNA library construction method described by Carninci P. et al. (2000), Genome Research 10(10):1617-1630 and directionally ligated into the pBluescript II SK (+) XR vector digested with XhoI (5' end) and BamHI (3'). Plasmid DNA was then transformed by electroporation into DH10B cells (Invitrogen) for propagation."

```
Query Match          3.3%;   Score 46;   DB 8;   Length 823;
Best Local Similarity  59.0%;   Pred. No. 0.16;
Matches   79;  Conservative   0;  Mismatches  55;  Indels   0;  Gaps   0;

Qy   515 AAGGCCATTGAGTAGGTTATGCTTGCCTTTGATTTTTTAATTGATGGATTGGTAAAAGTTTG 574
             ||||||||||| ||||||  |||| ||||| |||||| ||| ||||| |||||  ||||
Db   204 AGGGAAATCGTCTAGGTTATTCTTGTTATATTGATTTTAAATTTAAGCATCGAATAAAATGTTG 145

Qy   575 TTTTTTTTGATGATATATAATGCCATACCATCACCATCAAAATAAATAAAAATCTCCCGTT 634
             ||||  ||| || ||||| ||| || |||||||  || |||||||||||||| ||||
Db   144 TTTTGAGTATCAGCTTATGCCGAACCATTGGCATGATACACGCTTTTAAATTTGGGCCTT 85

Qy   635 TGGTGGCTAAGTAA 648
             ||| ||||| ||||
Db    84 TGTTTGAGTATTAA 71
```

Search completed: May 12, 2006, 11:35:43
Job time : 6046 secs

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 09:25:47 ; Search time 279 Seconds
                 (without alignments)
                 8792.240 Million cell updates/sec

Title: US-10-672-787-35_COPY_11357_12736
Perfect score: 1380
Sequence: 1 taaagcgtgaaccacagttaa..........acgacgcggcatattctttga 1380

Scoring table: IDENTITY NUC
                 Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues
Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
                 1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
                 2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
                 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
                 6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
                 8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
                 9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1380	100.0	96109	3	US-09-596-002-35	Sequence 35, Appl
2	1370.4	99.3	1410	3	US-09-540-236-673	Sequence 673, App
3	226.2	16.4	1416	3	US-09-328-352-1282	Sequence 1282, App
4	160.4	11.6	1830121	3	US-09-557-884-1	Sequence 1, Appli
5	160.4	11.6	1830121	3	US-09-643-990A-1	Sequence 1, Appli
6	160.4	11.6	1830121	3	US-10-158-865-1	Sequence 1, Appli
7	125.2	9.1	1323	3	US-09-543-681A-2830	Sequence 2830, Ap
8	117	8.5	1371	3	US-09-252-991A-7702	Sequence 7702, Ap
9	117	8.5	1401	3	US-09-252-991A-7861	Sequence 7861, Ap
10	117	8.5	1450	3	US-09-701-229-1	Sequence 1, Appli
11	100.6	7.3	1416	3	US-09-489-039A-5247	Sequence 5247, Ap
12	84.8	6.1	640681	3	US-09-790-988-1	Sequence 1, Appli
13	83	6.0	1368	3	US-09-107-532A-1574	Sequence 1574, Ap
14	69.8	5.1	567	3	US-09-252-991A-7787	Sequence 7787, Ap
15	67	4.9	2402	3	US-09-221-017B-881	Sequence 881, App
16	61.4	4.4	1372	3	US-09-530-836-1	Sequence 1, Appli
17	58.6	4.2	1242	3	US-09-147-928-1	Sequence 1, Appli
18	53.6	3.9	2423	3	US-08-714-918-86	Sequence 86, Appl
19	53.6	3.9	2423	3	US-09-265-315-86	Sequence 86, Appl
20	53.6	3.9	2423	3	US-09-265-315-86	Sequence 86, Appl
21	53.6	3.9	2423	3	US-09-266-417-86	Sequence 86, Appl
22	53.6	3.9	2423	3	US-09-528-709-86	Sequence 86, Appl
23	53.6	3.9	2423	3	US-09-527-745-86	Sequence 86, Appl
24	49.6	3.6	536	3	US-09-147-928-3	Sequence 3, Appli
c 25	49.4	3.6	699	3	US-09-134-000C-741	Sequence 741, App
c 26	48.6	3.5	1320	3	US-08-956-171E-80	Sequence 80, Appl
c 27	48.6	3.5	1320	3	US-08-781-986A-80	Sequence 80, Appl
c 28	48.2	3.5	1188	3	US-09-710-279-655	Sequence 655, App
c 29	48.2	3.5	3334	3	US-09-710-279-3574	Sequence 3574, Ap
c 30	47	3.4	1374	3	US-09-134-001C-1001	Sequence 1001, Ap
c 31	45.6	3.3	717	3	US-08-936-165A-78	Sequence 78, Appl
c 32	42.6	3.1	1368	3	US-09-902-540-9337	Sequence 9337, Ap
c 33	42.6	3.1	10486	3	US-09-902-540-997	Sequence 997, App
c 34	42	3.0	1407	3	US-09-530-838-1	Sequence 1, Appli
c 35	41.8	3.0	6693	3	US-08-961-527-195	Sequence 195, App
c 36	41	3.0	4010	3	US-09-710-279-3540	Sequence 3540, Ap
c 37	40.8	3.0	1141	3	US-09-806-708B-22	Sequence 22, Appl
c 38	40.8	3.0	1275	3	US-10-113-113-3	Sequence 3, Appli
c 39	40.2	2.9	1350	2	US-08-665-435A-1	Sequence 1, Appli
c 40	40.2	2.9	1350	2	US-08-665-435A-3	Sequence 3, Appli
c 41	40.2	2.9	1350	2	US-08-843-309-1	Sequence 1, Appli
c 42	40.2	2.9	1353	3	US-09-583-110-197	Sequence 197, Appl
c 43	40.2	2.9	1353	3	US-09-107-433-1773	Sequence 1773, App
c 44	40.2	2.9	1383	3	US-09-710-279-4356	Sequence 4356, Ap
c 45	39.4	2.9	3444	3		

ALIGNMENTS

RESULT 1
US-09-596-002-35
Sequence 35, Application US/09596002
Patent No. 6632636

GENERAL INFORMATION:
   APPLICANT: Lagace, Robert, E.
   APPLICANT: Patterson, Chandra
   APPLICANT: Berg, Kim, L.
   TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
   FILE REFERENCE: PM-0008-4 US
   CURRENT APPLICATION NUMBER: US/09/596,002
   CURRENT FILING DATE:
   PRIOR APPLICATION NUMBER: 60/140,121
   PRIOR FILING DATE: 1999-06-18
   NUMBER OF SEQ ID NOS: 41
   SEQ ID NO 35
   SOFTWARE: PERL Program
   LENGTH: 96109
   TYPE: DNA
   ORGANISM: M. catarrhalis
   FEATURE:
   NAME/KEY: misc feature
   OTHER INFORMATION: Incyte template ID No. 6632636 35
   PUBLICATION INFORMATION:
US-09-596-002-35

Query Match 100.0%; Score 1380; DB 3; Length 96109;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 TAAAGCGTGAACCACAGTTAACAAAACTTATCACCTCGCTCANTATNATTTTTAAATTGATC 60
          ||||||||||||||||||||||||||||||||||||||||| ||| ||||||||||||||||
Db  11357 TAAAGCGTGAACCACAGTTAACAAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 11416

Qy     61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCCTAGCCGTGCTTTGACTGGCTGCCAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11417 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCCTAGCCGTGCTTTGACTGGCTGCCAA 11476

Qy    121 TACTGCCCGATCAAGTAGTGTACCAACCTGATATATTTAGATGCCAAATGTGCCATCACCTTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11477 TACTGCCCGATCAAGTAGTGTACCAACCTGATATATTTAGATGCCAAATGTGCCATCACCTTT 11536

Qy    181 GAGTAAAATCTTGCTCAATTATCGGGGGCATCTTGACCGGATTAAATACACACTTATCCACCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11537 GAGTAAAATCTTGCTCAATTATCGGGGGCATCTTGACCGGATTAAATACACACTTATCCACCA 11596

Qy    241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCACCCACCCAAAAT 300
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11597 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 11656

Qy        301 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11657 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 11716

Qy        361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11717 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 11776

Qy        421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11777 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 11836

Qy        481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11837 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 11896

Qy        541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCCATACC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11897 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCCATACC 11956

Qy        601 ATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTTGTTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11957 ATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTTGTTGT 12016

Qy        661 GATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12017 GATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 12076

Qy        721 GGCATCGTCTTGGTAGATGATGGCGGTTTGGCAATTTTCAAAAATACGCAATTTTTGTGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12077 GGCATCGTCTTGGTAGATGATGGCGGTTTGGCAATTTTCAAAAATACGCAATTTTTGTGC 12136

Qy        781 CAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12137 CAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTGT 12196

Qy        841 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12197 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCAC 12256

Qy        901 CACCAAATCCATATTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTACCGCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12257 CACCAAATCCATATTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTACCGCC 12316

Qy        961 GACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGCT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12317 GACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGCT 12376

Qy       1021 TTTGGCATTTGAGCCTGTGATGGCGATGATTGGTGTGTTTGTTGCTTGAGTTAGATCTCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12377 TTTGGCATTTGAGCCTGTGATGGCGATGATTGGTGTGTTTGTTGCTTGAGTTAGATCTCG 12436

Qy       1081 TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTTT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12437 TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTTT 12496

Qy       1141 GGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12497 GGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGGC 12556

Qy       1201 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTTT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12557 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTTT 12616

Qy       1261 ATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      12617 ATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGAC 12676

Qy       1321 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCATATTTTTTTGA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      12677 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCATATTTTTTTGA 12736

RESULT 2
US-09-540-236-673/c
; Sequence 673, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
;  TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 2709.2005-001
;  CURRENT APPLICATION NUMBER: US/09/540,236
;  CURRENT FILING DATE: 2000-04-04
;  NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 673
;   LENGTH: 1410
;   TYPE: DNA
;   ORGANISM: M.catarrhalis
US-09-540-236-673

  Query Match             99.3%;  Score 1370.4;  DB 3;  Length 1410;
  Best Local Similarity   99.6%;  Pred. No. 0;
  Matches 1374;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1392 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 1333

Qy         61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 120
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db       1332 CATGCTAGCACAAGCAGGCGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 1273

Qy        121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTAGATGCCAAATGTGCATCACCTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1272 TACTGCCCGATCAAGTGTACCAACCTGATATATTTAGATGCCAAATGTGCATCACCTTT 1213

Qy        181 GAGTAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACTATCCACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1212 GAGTAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACTATCCACCCA 1153

Qy        241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1152 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 1093

Qy        301 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1092 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 1033

Qy        361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db       1032 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCCATCAATATCATCAATATA 973

Qy        421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        972 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 913

Qy        481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        912 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 853

Qy        541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCCATACC 600
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db        852 TTTGATTTTTAATTGATGGATTGGTAAAAGCTTGTTTTTTTGATGATATAATGCCATACC 793

Qy        601 ATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTTGTTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        792 ATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTTGTTGT 733

Qy        661 GATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        732 GATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 673
```

```
RESULT 3
US-09-328-352-1282/c
Sequence 1282, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
  APPLICANT: Gary L. Breton et al.
  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
  FILE REFERENCE: GTC99-03PA
  CURRENT APPLICATION NUMBER: US/09/328,352
  CURRENT FILING DATE: 1999-06-04
  NUMBER OF SEQ ID NOS: 8252
  SEQ ID NO 1282
  LENGTH: 1416
  TYPE: DNA
  ORGANISM: Acinetobacter baumannii
US-09-328-352-1282

Query Match    16.4%;   Score 226.2;   DB 3;   Length 1416;
Best Local Similarity  52.1%;   Pred No. 1.1e-58;
Matches 718;   Conservative 0;   Mismatches 593;   Indels 66;   Gaps 7;
```

(Nucleotide alignment between Query (Qy) and Database (Db) sequences follows, paired with vertical match bars.)

```
Qy     1196 TCGGCTGAGCTTAAGAGTTCACTGTCCAAAGCTGCCAAAATGACAATGAACGCCTGTGGGC 1255
            ||||||||||  ||    |    ||   || ||| |||    || ||| |   ||| | ||
Db      269 TCTGCTTGTAATAATAATTCTTGATCAAGCTGACCAAAAACTGGTTTTAACACAGCCCGGA 210

Qy     1256 AGTTTATCTGCCCAAGGTAGGAATTGGGGATTACCATCGATGAGATGCTAACTTTGTGACCGGTGA 1315
            |  | ||| || | ||| |  |    ||||  ||| |||||  |   || |||  |||||
Db      209 ATCTGATCGTGTCCGGGAGGTGTGGGGCGGAATCCGTTACAGCAAAACTGGTTACCCTTGT 150

Qy     1316 TTGACCAAAAAAATTCACAGCTGACAGTCGCCAAAACCGACGACGGCCATAT 1372
            |||| |||||| |||||||||| |  ||| | ||||| | |   |  ||||
Db      149 TCATGCAGAAAATTTACAGCAGAAACACCTGATATTCCCAAGCCTGCTACAACTTTT 93
```

RESULT 4

; US-09-557-884-1/c
; Sequence 1, Application US/09557884
; Patent No. 6506581
;   GENERAL INFORMATION:
;   APPLICANT: Fleischmann et al.
;   TITLE OF INVENTION: The Nucleotide sequence of the Haemophilus influenzae Rd Genome, Fragments
;   Thereof, and Uses Thereof
;   NUMBER OF SEQUENCES: 1
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville
;   STATE: MD
;   COUNTRY: USA
;   ZIP: 20850
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3 1/2 inch diskette
;   COMPUTER: Dell Pentium
;   OPERATING SYSTEM: MS DOS v6.22
;   SOFTWARE: ASCII Text
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/557,884
;   FILING DATE: 25-Apr-2000
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/476,102
;   FILING DATE: JUN-5-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Michelle S. Marks
;   REGISTRATION NUMBER: 4,971
;   REFERENCE/DOCKET NUMBER: PB186P3
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 301-309-8504
;   TELEFAX: 301-309-8439
;   INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1830121 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-557-884-1

Query Match   11.6%; Score 160.4; DB 3; Length 1830121;
Best Local Similarity 49.6%; Pred. No. 9.2e-37;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```
Qy        1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTTCGCTCCATTATATATTTAAATTGATC 60
Db  1203428 ------------GAATAATTCAATATCGCCGATTAATATCCCACTCCGCTTTAA 1203388

Qy       61 CATGCTAGCACAGGTGCACAAAAGCACAGCCCTAGCCGGTGTTGCTTGGCTGCCAA 120
Db  1203387 GTGCGGGTTTGAATTTCTGGTGTTTTACCGGAGCCCTGATGATGATGATCGC 1203328

Qy      121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGTGCCAAAATGTGCATCACCTTT 180
Db  1203327 TTTCAAGTAACCATTCCCTGATTAAAACTACCAGTATGAAGAGGGATTATTTGAGGAAGTT 1203268

Db  1204331 TGTTGGGGCGTGAAAATTCTATCGCTGCTTGTTGTTCCATTGTATCGAACAA------ATAACTTTG 1204278
Qy      181 GAGTAAAATCTTGCTCAATTATCGGGGCCATCCTTGACCGGATTAAATACACACTATCCACCCA 240

Db  1204277 CGATGAAAAATTTTGCAACAAGCAGCGCACCAACCATCTCGACCAAAATAAAGTGTGG 1204218
Qy      241 CTGCATTAACAATATTGCTGACAGCTCACTAGAAAATTCTTGAACAAAAATAACAAATTAAAGTGTGG 300

Db  1204217 TTGATTAAATTAATTAGCTAATTCTTGAAAAAATCAGCCCCTTTTCCGTCTCCGCCAACCCAAAAAT 1204158
Qy      301 CAATGCCAGTGATGATAATGCCATAAAACCGTGCCAAGCCCATCAACAATTGCTGCAATGGTTGA 360

Db  1204157 CAAAATGCAATTTACAGCAGAAACACCTGATATTCCCAAGCCTGCTACAACTTTT 1204107
Qy      361 TTTTCGGCATTTTGAGCTGTAGATGGGATGATTGGTGTGTTCTTGGTTGGCTTAGATCTC 1079

Db  1203987 TTGAAAACGATGATCTAACCCTTTGAAAAGTTGAAAATGACGAAGTGCGGTACAGAAGTGAATTGAAATCTAAAATT 1203867
Qy      421 TTCACAGCGATGGCGGGTAGCCCTTTGAAAGTTCATTAATCAAGAAATATGCCTTGCTACTAAATACATTCCATGGG 480

Db  1203866 GCCCATTTTCAGTTTTTAGCCAATAATCCGGACTAATTTCCGCAAAAGAAAGGGGTATGTTT 1203807
Qy      481 AAGATTGGGCAGCTGCTGTCCTAAGGCAAGGTGCAAGTGGCATTGAGTGACGTAAGCTATTGTTG 540

Db  1203807 GCCATTTTCAGTTTTTAGCCAATAATCCGGACTAATTTCCGCAAAAGAAAGGGGTATGTTT 1203807
Qy      541 TTTGATTTTTAATTGATGGATTGGTGAAGAAGTTTGTTTTTTTTTGGGTGTGGCTAAGTAAGCTATTGTTG 600

Db  1203807 AACCAAATGTAGCTTCTTCACAAGGTAAAATCACTTCATCTTTTACCATTAATAATAATGCCATACC 1203718
Qy      601 ATCACCATCATCAAATAAATAAAATCTGC--CGTTTGGGTGGCTAAGTAGTAAGCTATTGTTG 659

Db  1203718 --TATCTTCATTGTTCAACACCTACTTTAGCATTTCATTAGATAATAGCGTAATTTGCTT 1203658
Qy      660 TGATAATAGGTTGTGATTTGGGTGCCACAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGGCTAAAAT 719

Db  1203658 GGGCGATAATCTTCTAAATCCATATAGCATTCCATATGACATCAGTCAGTGGAGAGATTCAAGACAG 1203598
Qy      720 TGGCATCGTCTTGGTAGATGATGCGGGTTTGGCAATTTCAAAAATATACGGCAAATTTTGTG 779

Db  1203597 CATAAAGTTCACAAACTTCATTC---AACAATGACAAAGGGGGAATCCCAATATTTCCGC 1203541
Qy      780 CCAAATAATCCACCATGCCGATCAAGTCGCTGCGATGAGAGATAGCGAGAGATTAAAATTG 839

Db  1203540 CGACACCAACAATCATGCCTGCATGTTTTGCCCATCTCGCCTACTAAGTAAGTTACGGTAC 1203481
Qy      840 TTGCCCCTTGAGCTGCCAAATTTGAGATATGCCTCAAAGCTAAAACTTGACAGCTCAAGCA 899

Db  1203480 TTTTTACCATTTGAACCTGTAATCCCACAATTGGTCTTCGCGTCGCGGGCA------ 1203429
Qy      900 CCACCAAAATCCATATTTTGGATCGTTCAAATAATTCAAGTGCAGGCGGTGCCAATATTACCGC 959

Db  1203597 TCGCTGCCGAGCGCTTTAAGCTATAAAGTTGTCTCAAAGCTGAAAAAGCTCTAGTA 1203598
Qy      960 CCATCACACTTCACACCAGCAGCTTCGCCATGTGTTTCGCCCTACTTTTCATAAACTAAAGTTACGGTAC 1019

Db  1203387 GAGACTTGCACAGCGGGGACAATAATACCA 1203268
Qy     1020 TTTTCGGCATTTTGAGCTGTAGATGGGATGATTGGTGTGTTGTTCTTGGTTGGCTTAGATCTC 1079

Qy     1080 GTGCTTTGAGTGTATCAATGAACAATGAACATCGCTGATCACAGGAATGCCCTTGAGCTT 1139
Db  1203480 TTTTACCATTTGAACCTGTAATCCCACAATTGGTCTTCGCGTCGCGGGCA------ 1203429

Qy     1140 TGGCCGGCGACAAATACTCGGCGGTTCTTGGGGTTAATCGCGGCGATGATGATGATCGC 1199
Db  1203387 GTGCGGGTTTGAATTTCTGGTGTTTTACCGGAGCCCTGATGATGATGATCGC 1203328

Qy     1200 CTGAGCTTAAGAGTTCACTGTCCAAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTT 1259
Db  1203327 TTTCAAGTAACCATTCCCTGATTAAAACTACCAGTATGAAGAGGGATTATTTGAGGAAGTT 1203268
```

RESULT 5
US-09-643-990A-1/c
; Sequence 1, Application US/09643990A
; Patent No. 6528289
; GENERAL INFORMATION:
; APPLICANT: Robert D. Fleischmann
;   Mark D. Adams
;   Owen White
;   Hamilton O. Smith
;   J. Craig Venter
; TITLE OF INVENTION: The Nucleotide sequence of
;   the Haemophilus influenzae Rd Genome, Fragments
;   Thereof, and Uses Thereof
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Human Genome Sciences, Inc.
;   STREET: 9410 Key West Avenue
;   CITY: Rockville,
;   STATE: MD
;   COUNTRY: USA
;   ZIP: 20850
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3 1/2 inch diskette
;   COMPUTER: Dell Pentium
;   OPERATING SYSTEM: MS DOS v6.22
;   SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/643,990A
;   FILING DATE: 23-Aug-2000
;   CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/487,429
;   FILING DATE: 1995-06-07
;   APPLICATION NUMBER: 08/426,787
;   FILING DATE: 1995-04-21
; ATTORNEY/AGENT INFORMATION:
;   NAME: Kenley K. Hoover
;   REGISTRATION NUMBER: 40,302
;   REFERENCE/DOCKET NUMBER: PB16P1C1
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 301-610-5790
;   TELEFAX: 310-309-8439
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 1830121 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-643-990A-1

Query Match   11.6%;   Score 160.4; DB 3; Length 1830121;
Best Local Similarity   49.6%;   Pred. No. 9.2e-37;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

Qy   1 TAAAGCGTGAACCATGTTAACAACTTATCACCTCGCTCATTATATAATTTTTAAATTGATC 60
Db 1204439 TAAACATTGAGCTAAATGCCTAAATTCTTCGCCCGGCCTTTTCAAAAGAAGCAAACTGATC 1204380

Qy   61 CATGCTTAGCACAAGCAGGGTGACAAAAGCACAGCGCCTAGCGCGTGCTTTGACTGGCTCC 120
Db 1204379 GAGACTTGCACAAGCAGGCGGCCGACAATAATACCA---------------------- 1204332

Qy   121 TACTGCCCGATCCAAGTGTACCAAACCTGGATATATTTTTAGATGCCAAAATGTGCATCAC 180
Db 1204331 TGTTGGGCGTAAAAAATTCTATCGCTTGTTCCATTGTATCGAACAA---- 1204278

Qy   181 GAGTAAATCTTGCTCAATTATGGGGGGCATCTTGAGCGGATTAAATACACTATCACCCA 240
Db 1204277 CGATGAAAAATTTTGCAAGCACGGCAAACCATCTCGACCAAAACAATACAGAAATGTGTGG 1204218

Qy   241 CTGCATAACATATGCTGACAGCTCACTCAAAAATCTTGACCTTTACCAAGCCCACCCAAAT 300
Db 1204217 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCGCCCTTTTCCGTCTCCGCCTAGCA 1204158

Qy   301 CAATGCCAGTGATGATATGCCATAAAACCGTGCCCAAGCCCATCAAATTGCTCGTGCAATG 360
Db 1204157 CAAATGCAATTTACCCCTCCAATATA--------AAGCCCAGCCAATGCCAGCAACTGTACT 1204107

Qy   361 GCCGATGTTTGTGTACCTTTGAATCATTAAAAATATGCCTTGCTATCATATCATCAATATA 420
Db 1204106 CCCCACATTTGTTGTCTTTAGAGTCATTATTAATCCAACATTAGCTTGCACTAA 1204047

Qy   421 TTCACGCGATCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 480
Db 1204046 TTGAAAACGATGAATCTAACCCTTGAAATGACGAAGTGCGGGTACGAATTGAATCTAAATT 1203987

Qy   481 AAGAGTTGGCAGCTGTGCCTAAGGCAGTGCAGATAAGGCATTGAGTGAGGTTATGCTTGCC 540
Db 1203986 AATACCTATAGTTGTGTCCCAAATGCTGTTGCTGCCAAAATGTTCATAATTATGGCAACC 1203927

Qy   541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTGTTTTTTTTGATGATATAATGCCATACC 600
Db 1203926 AACCAATGTAGCTTCTTCACAAGGTAAAATCACTTCATCTTTTTACCATTAAAATAATTGCTT 1203867

Qy   601 ATCACCATCATCAAATAAAATAAAAAAATCTGC-CGTTTGGTGGCTTAAGTAAGCTATTTGTG 659
Db 1203866 GCCCATTTTCAGTTTGTTGCCAATAATCCGCACTATTTTTCCGGAAAAGAAGACGGTATGTTT 1203807

Qy   660 TGATAATGGTGTGATTGGGGGGCAAGCCTTGTCAGTGGTGGTCAAGCATTGCTTGGCTAAAT 719
Db 1203806 CGCTTGATTTTTCGTTTTCCCCAAAAGTCAGCC-- 1203775

Qy   720 TGGCATCGTCTCTTGGTAGGATGGGCGGTTGGCAATTTCAAAAAATACGCAATTTTTGTG 779
Db 1203774 ---TATCTTCATTGTTCAACACACTACTTTAGCTAATTGTAGATAAAGCGTTAAATTTTGCTT 1203718

Qy   780 CAAAATAATCCACCATCGTGCCGATCAAGATGATGCAGCGGGAGAGATTTAAAATTG 839
Db 1203717 GGCGATAAATCTTCTAAATCCATATACGGATCCATATGATCTTCAGTCACGGTTCAAGACAG 1203658

Qy   840 TTGCCCCTTGAGCGTGCCAAATTTGAGGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA 899
Db 1203657 TCGCTCGCCGCAGCTTTTAAGCTATAAGTTGTCTCAAGCTGAAAAGCTCTAGTA 1203598

Qy   900 CCACCAAATCCATATATTTGGGATCGTAAATAATCAGTGCAGGCGGTGCCAATATTACCGC 959
Db 1203597 CATAAAGTTCACACAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAATATTTCCGC 1203541

Qy   960 CGACACCAACAATCATGCCTCACTGTTTGCCATCTCGCCTACTATGTCGTACCAGTGC 1019
Db 1203540 CCATACCAACTTTCACCACCAGCCTTTCGCCCATTCCATAAAACTAAAGTAGTTACGGTAC 1203481

Qy   1020 TTTTTGGCATTTGAGCCTGTGATGGGGATGGAATGGTTGGTGTGGTTGTTGCTTGAGTTAGATCTC 1079
Db 1203480 TTTTTACCATTTGAACCTGTAATCCCCACAAATGGCTTTGTCGCCTGCGGGGCA------ 1203429

Qy   1080 GTGCTTTGAGTGTATCAAATGAACAATTGAACATCGTGCTTCACAGAAATGCCTTGAGCTT 1139
Db 1203428 ------GAATAATTCAATATCGCGGGTTGTTAATGCGGCCTTGTCGATGATCGGTCGG 1203388

Qy   1140 TGGGCGCGACAATACTCGGGCGTTCTTGGGGTTTAAATGCCCGATGATCGTGATCCGGCTTAA 1199
Db 1203387 GTGCGGCGGTTTGAATTTCTGGTCGTTTTTACCGGCAAGCCCCTGCAATAACAATCATATCGC 1203328

Qy   1200 CTGAGCTTAAGAGTTCACTGTTCCAAGCTGCCAAAATGACAATGAAGCGCTGGGGGCAGTT 1259

```
Db  1203327  TTTCAAGTAACCATTCCTGATTTAAACTACCAGTATGAAGAGGGATATTTGAGGAAGTT  1203268
                        ||||        ||  ||||||  ||||||||||||||||  || ||||
Qy     1260  TATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAAACTTTGTGACCGTGATTGA  1319

Db  1203267  TATCAAATACCAGTAGGATTTTTTCGAGTATCAATCACACGAATATTTAGCCTGTTGGGATA  1203208
                        |||||||  ||||        || |||| |||  || ||||||||  |
Qy     1320  CCAAAAAATTCACACGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGGCATATTTTTTT  1377

Db  1203207  AGAGATAATCCACACAGAAGAAAGACCTGTTTTTGCCAAGCCCGATGATGATAGTAATATTTTTT  1203150
```

```
RESULT 6
US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
;  APPLICANT: Fleischmann et al.
;  TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
;  Patent No. 6846651
;  TITLE OF INVENTION: Thereof, and Uses Thereof
;  FILE REFERENCE: PB186P2CID1
;  CURRENT APPLICATION NUMBER: US/10/158,865
;  CURRENT FILING DATE: 2002-06-03
;  PRIOR APPLICATION NUMBER: US 09/557,884
;  PRIOR FILING DATE: 2000-04-25
;  PRIOR APPLICATION NUMBER: US 08/476,102
;  PRIOR FILING DATE: 1995-06-07
;  PRIOR APPLICATION NUMBER: US 08/426,787
;  PRIOR FILING DATE: 1995-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 1
;  LENGTH: 1830121
;  TYPE: DNA
;  ORGANISM: Haemophilus influenzae
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (4747)..(4747)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (9921)..(9921)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (10150)..(10150)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (29298)..(29298)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (36543)..(36543)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (36551)..(36551)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (36636)..(36636)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (40808)..(40810)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (44416)..(44416)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
```

```
;  NAME/KEY: misc_feature
;  LOCATION: (44905)..(44905)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (44975)..(44975)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (45593)..(45593)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (45732)..(45732)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (47036)..(47036)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (51334)..(51334)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (51602)..(51602)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (51786)..(51786)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (51805)..(51805)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (55369)..(55369)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (65309)..(65309)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (65313)..(65313)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (80024)..(80024)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (100091)..(100091)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (102696)..(102696)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (102748)..(102748)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (105121)..(105121)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (117136)..(117136)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139510)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
```

```
Qy        1  TAAAGCCTGAACCATGTTAACAAAACTTATCACTTCGTCATTATTAAATTTGATC      60
             ||||||||
Db  1204439  TAAACATTGAGCTAAATCCGTAAATTCTTCGCGCCGGCGGCTTTTCAAAGAAGCAAACTGATC  1204380

Qy       61  GAGTAAATCTTGCTCAATTATTGGGGCATCTTGACCGATTAAATACACACTATCCACCCA     120
Db  1204379  GAGACTTGCACAAGCAGGCGACAAATAATACCA------------TATCTCCGCGTTTGCAA  1204332

Qy      121  TACTGCCCGGATCAAAGTGTTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACTTT    180
Db  1204331  TGTTGGGGCGGTAAAAAATTCTTATCGCTTGTTCCATTGTATCGGACAA------ATAACTTTG  1204278

Qy      181  GAGTAAATCTTGCTCAATTATTGGGGCATCTTGACCGATTAAATACACACTATCCACCCA     240
Db  1204277  CCCCACATTTGTTGCTTTAGAGTCATTAATCCAACGAAATGCCATTAGCTTGATGCACTAA  1204218

Qy      241  CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT    300
Db  1204217  TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTCCGGTCTCCGCGTTAGCAA  1204158

Qy      301  CAATGCCAGTGATGATGATATGCCAATAAACCGTGCCAAGCCCATCAATTGCTGCAATATA    360
Db  1204157  CAAATGCAAATTTACCCTCAAATATA-------AAGCCCAGCCGAAATGCAGCGAACTGTACT  1204107

Qy      361  GCCGATGTGTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA    420
Db  1204106  CCCCACATTTGTTGCTTTAGAGTCATTAATCCAACGAAATGCCATTAGCTTGACTAA  1204047

Qy      421  TTCACAGGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATCCATGGG    480
Db  1204046  TTGAAAACGATGATCTAACCCTTTGAAATGACGAAGTGCGGTAACTAAATT  1203987

Qy      481  AAGATTCGGCAGCTGTGCCTAGGCAACAGTGCAGATAAGCATTGAGTAGGTTATGCTTGCC    540
Db  1203986  AATACCTTATAGCTTGTGCCAATGCTGTTGCTCGCCAAAAGTTCATATATAATTATTGGCGACC  1203927

Qy      541  TTTGATTTTTAATTGATGGATTGGTAAAAAGTTTGTTTTTTTGATGATAATAATGCCATACC    600
Db  1203926  AACCAAATGTAGCTTCTTCACAAGGTAAAATCACTCATCTTTTTACCATTAAAATATTGCTT  1203867

Qy      601  ATCACCATCATCAAATAAATAAAAAATCTGC-CGTTTGGTGGGCTAAGTAAGCTATTGTTG    659
Db  1203866  GCCATTTTCAGTTTTTAGCCAATAATCCGACTAATTTTTCCGGAAAAGAAGCAACGGTAGTTTT  1203807

Qy      661  TGATAATGGTGTGATTTGGGGCAACCGCTTGTCAGTGGTGGTCAAGCAATTGCTTGGCTAAAT    719
Db  1203806  CGCTTGATTTTGCGTTTTCCCCAAAAAGTCAGCC-----------------------  1203775

Qy      720  TGGCATCGTCTCTTGGTAGATGATGGTTGGCAATTTCAAAAATACGCAATTTTTGTG    779
Db  1203774  ---TATCTTCATTGTTCAACACACTACTTTAGCATTATGATTAAATCGTATTTTGCTT  1203718

Qy      780  CCAAAATAATCCACCATGCCCGATCAGTGGCCGGACTCAAAGATGAGATGAGATTTAAAATTG    839
Db  1203717  GGCGGATAAATCTTCTAAATCCATATAGGGATCCATATATGATCTTCAGTCACGGTTCAAGACAG  1203658

Qy      840  TTGCCCCTTGAGCTGCGCCAAATTTGACATATGCTCAAGCTGAAAACTTGACAGCTCAAGCA    899
Db  1203657  TCGCTGCGCGCAGCCTTTTAAGCTTATAAGTTGTGTCTCAAAACTAGAAAAGCCTAGTA  1203598

Qy      900  CCACCAAATCCATATTTGGATCGTTAATTAATTCAAGTGCAGGCGGTGCCCAATATTACCGC    959
Db  1203597  CATAAAGTTCACAATCTTCATTC----AACAATGACAAAGCGGGAAACCCAATATTTCCGC  1203541

Qy      960  CGACACCAACAATCATGCCTGCACATGTTTTGCCATCTCGCCTTACTAATGTCGTAAACAGTGC    1019
Db  1203540  CCATACCAACTTTCACCACGACAGCTTTCCGCATTCATAAATCATAAGTAGTTACCGGTAC  1203481

Qy     1020  TTTTGGCATTTGAGCCTTGTGATGGCGGATGATGGTGTGTTTGCTTGAGCTAGATCTC    1079
Db  1203480  TTTTACCATTGAAACCTGTAATCCCCACAATTGGGCTTGCTTTGTCGCGTTGCGCGGGCA------  1203429

Qy     1080  GTGCTTTTGAGTGTATCAAATGAAACAATTGAACATTCGCTGATCACAGGAGGAAATGCCTTGAGCTT    1139
```

Query Match 11.6%; Score 160.4; DB 3; Length 1830121;
Best Local Similarity 49.6%; Pred. No. 9.2e-37;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```
Db 1203428 -----------------GAATAATTCAATATCGGCGATTACTTCCACTCCCGCTTTAA 1203388
                            |||| |||| ||  |||   |||
Qy    1140 TGGCGGCGACAATACTCGGCGGTTCTTGGGGTTAATGCCGGGACCTGATGATGATCGGACGG 1199
                  |||  || | |||     |    || |||  |||  |
Db 1203387 GTGCCGGTTTGAAATTTCTCGGTGTTTTTACCGCCAAGCCCTGATAACAATCATATCGC 1203328
                  || |||  ||||   |||    ||  |  |
Qy    1200 CTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAAATGAAACGCCTGTGGGCAGTT 1259
                  |||     ||| | || |||   ||     ||
Db 1203327 TTTCAAGTAACCATTCCTCGATTTAAACTACCAGTAATGAGAGGGATATTTTGAGGAAGTT 1203268
                  |||  |  ||| || || || |||   ||
Qy    1260 TATCTGCCAAGGTAGGATTGGGATTTACCACCATCGATCAGTGCTAAACTTGTGACCGTGATTGA 1319
                  |||  ||| || || ||| ||   |||||
Db 1203267 TATCAATACCAGTAGGATTTTTCGGAGTATCAATCACACGAATATTAGCCTGTGTTGGGGATA 1203208
                  |||||||  |  || || |||
Qy    1320 CCAAAAAAATTCACAGCTGACAGTTCCAGAGAGCTGCCCAAAACCGACGACGGGCCATATTTTTT 1377
                  ||||||  ||  |  ||  |||   |||||
Db 1203207 AGAGATAATCCACACAAGAAAGACCTGTTTTTGCCAAGCCCGATGATAGTAATATATTTT 1203150
```

```
  Query Match             9.1%;  Score 125.2;  DB 3;  Length 1323;
  Best Local Similarity   53.5%;             Pred. No. 1.1e-27;
  Matches  352;  Conservative    0;  Mismatches  273;  Indels  33;  Gaps  3;

Qy    723 CATCGTCTTGGTAGATGATGGCGGTTTGGCGCAATTTTTCAAAAAATACGGCAATTTTGTGTGCCA 782
Db    658 CATCACTCTCGGTTATAGACACAGGTTTTAGCATTATATCATAAATACGTAATTAGTTAGCTTCAC 599
Qy    783 AATAATCCACCCATGCC---ATCGTGCCGATCAAGATGATCAGCGGAGAGGAGGATTTAAAATTG 839
Db    598 GGTATATTGTTCTAATCCTAACGGATAGCGATTCATCATGATCTTCGCTAATAATTAATACCG 539
Qy    840 TTGCCCCTTGAGCTGCCAAATTTGAGATAATGCTCAAGCTGAAAACTTGACAGCTCAAGCA 899
Db    538 TGGCGAGCCCGCGTCGTTAAGCTATAGGTTGTGTCTCTAAATTGAAAGTCTGATAGCTTCTAAGA 479
Qy    900 CCACCAAATCCATATTTTGGATCCTTTAAATCAAGTGCCAGGGGTGCCAATAATTGCCGG 959
Db    478 CATATAAGTCATAAGGCTGT---GTTAATAACGTTAAAGCAGGTACACCAATAATTGCCAC 422
Qy    960 CGACACCAACAATCATGCCTTGCAATGTTTTGCCATCTCGCCTACTAAATGTCGTAACAGTGC 1019
Db    421 CGACACCGACAGGGAGGATCCCGTCAGCTTTAGCCATTTCCCCACACGTCGTTACGGGTAC 362
Qy   1020 TTTTGGGCATTGAGCCTGTGATGGGCGGATGAGTTGGGTGTGTTGTGCTTGAGTTAGATCTC 1079
Db    361 TTTTGCCATTAGAGCCTGTTATCGGAACAATAGAGTGCTTTTGCCTCACGG------- 312
Qy   1080 GTGCTTGAGTGTTCATAATGAACAAATTGAACATCGCTGATTCACAGGAATGCCCTTGAGCCC 1139
Db    311 ------------CAAAATAGCTCAATATATCACCGACAAATTTCCACTCCATGAATCTC 269
Qy   1140 TGGCGGCGACAATACTCGGCGGTTCTTGGGGTTAATGCCGGCTGATGATGATCGGACGG 1199
Db    311 ------------CAAAATAGCTCAATATATCACCGACAAATTTCCACTCCATGAATCTC 269
```

```
  Query Match             8.5%;  Score 117;  DB 3;  Length 1371;
  Best Local Similarity   48.4%;             Pred. No. 3.5e-25;
  Matches  506;  Conservative    0;  Mismatches  500;  Indels  39;  Gaps  5;

Qy    17 TTAACAAAACTTATCACCTCGTCGCTCTCATTATAATTTAAATTGATCCATTGCCTAGCACAGCA 76
Db   1349 TTGGGGAACAGGCGTCCGGCGTTCTTCGAAGTTCTTCGAACATGTTCCAGGCTTCGGCAGGCC 1290
Qy    77 GGTGACAAAAGCAGCAGCCCTAGGCGGTGCTTGACTGGCTGCCAATACTGCCGCTCAAGT 136
Db   1289 GGCGAACAAGCACCACCCGGATGGAGCCATCGGCCTTCGC-------GGGCGGCCTTCGCCGG 1239
Qy   137 GTACCAACCTGATATATTTTAGATCGCCAAAATGTGATCACCTTTGAGTAAATCTTGCTCA 196
Db   1238 ACTGCTTCGTCGTCCAGCGTTGCCACGGCCACCGGTACCGCGGTTGCCCAGTGCCTGGGCA 1179
Qy   197 ATTATCGGGGGCATCTTGACCGGATTAAAATAACACATCCACCACTGCCATAACATATGCT 256
Db   1178 ATCAGCCCCGGCGGTCACGGCCAAGCAGTACCACGCGCGGCGGGCCGAGACCGGCTCG 1119
Qy   257 GACAGCTCACTAAAAATCTTGACCTTTACCAAGCCACCACCGCCGGCGAAAATACTGCCAGTGATGAT 316
Db   1058 TCGATGT---CGGCACCCAGCCCCTGATCGCCGCCAGGCGCCGACGGGCTTGGTGGCC 1002
Qy   317 ATGCCATAACCGCCGCCAGCCCATCCAAATGCTGAAATGGTTGACCGCGATGTTTGTACCT 376
Db   1001 TTGGAAATCGTCGTAGTAGCTCACGCCCATCGCGTCCGGTACCCACTGCGTCCCAGCC 942
Qy   377 TTTGAATCATTAAAAATAGCCTTGCTCTATCAATATCATCAATATATTCACAGCATGATCCGGT 436
Db    941 AGGCCGGAAAACGCCTTCAGCGGGCGCGAGCATCCATGGGGAAGATTGGCAGCTGTG 882
Qy   437 AGCCCTTTGAAAGTCTTAAGGTATCAAGCATGGCATCCATGCCGATAGTTGTTGTACCT 496
```

```
Qy   497 CCTTAGGCAAGTGCAGATAAGGCAATTGAGTGAGTAGGGTTATGCTTGCCTTTGATTTTTAATTGA 556
         ||||||||
Db   881 CCCAGCGCCAGCGCCGACGGCGTTGGAATAGTTGTGGGCGCCCACGGATCTTCAGTTCG 822

Qy   557 TGGATTGGTAAAAGTTTAGTTGTTTTTTGATGATATAATGCCATACCATCACCATCAAAT 616
         ||||||
Db   821 CCAACGCGAGCGCAGCTTGTCGACTGGAACGCCAGCACTTCTGGCCGTCTTCCTCGATC 762

Qy   617 AAATAAAAATCTGCGCGTTTGGTGGCTAAGTAAGCTATTTGTTGTGATAATGGTGTGATTT 676
         ||||
Db   761 AGGCGCGAAAGC------------CTTGAAGTCCGGCGGCTTGTTGTTCAGGCCGAA 724

Qy   677 GGGGCAACCGCTTGTCGTCGTGTGGGTCAGTCATTGCATCGTCTCTTGGGTAG 736
         ||||
Db   723 CGACCAGCACGGACGGTACGGCCGATCACGCGGTCA--GGGCATCGGCGCGAATTC 666

Qy   737 ATGATGGCGGTTTGGCAATTTTCAAAAAATAGGCAATTTGTGCCAAATAAATTCACCACATG 796
         ||||
Db   665 ACCACGACCTGGCGCGGGAGAGATCCGGTGGATAGTTCACAGGATGCGGGT 606

Qy   797 CCATCGTGCCGATCAAAGATGATCACGCGGAGAGATTTAAAATTGTTGCCCCTTGAGCTGCC 856
         ||||
Db   605 CCGTCGTAGCGGATCCATATGGTCTTCGCTGAGCAGTGCGTCGCCACTGGGA 546

Qy   857 AAAATTTGAGATATGCTCAAGCAGCTGAAAAACTTGACAGCTCAAGCAGCACCACCAAAATCCATATTT 916
         ||||
Db   545 AGGGCGATCGCAGGGTTTCCACCTGGAAGCTCCAACACGCTACACTCGAGTGTCG 486

Qy   917 TGGATCGTTAATAATTCAAGTAGTGGCAGGGCGGTGGCCAATATTACCGGCCCGGACCAACAATCATG 976
         ||||
Db   485 TCG---GCCAGCAGTCGGAGCGCCCGGGGTTGCCGCGGGGACACGCTTG 429

Qy   977 CCTGCATGTTTTTGCCATCTCGGCCTACTAACGTGCGTTTTTGGCATTTGAGCCT 1036
         ||||
Db   428 TCCGGCGGCCCACCGCCCATTTCGCCCAGCGGTGGCTCACGGTTGCTCTTCGGCGTTGGAACCG 369

Qy  1037 GTGATGGGCCGATGAATTGGGTGTGGTTTG 1061
         ||||
Db   368 GTGATGGGCCACGACTCGGGGCCTTCG 344
```

RESULT 9
; Sequence 7861, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7861
; LENGTH: 1401
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa

Query Match   8.5%; Score 117; DB 3; Length 1401;
Best Local Similarity   48.4%; Pred. No. 3.6e-25;
Matches 506; Conservative 0; Mismatches 500; Indels 39; Gaps 5;

```
Db    74 GCGGACAAACAGCACCGGACATCGGCGCCTTCGC-----------GGGCCAGCTCGGGCGGCGGCTCCGG 124

Db   137 GTACCAACCTGATATATATTTTAGATGCAAAGTGCATCACCTTTGAGTGTAAATCTTGCTCA 196

Db   125 ACTTGCTTCGTCGTCAGCGGTTGCGACGGCGCCACCAGCACCGGCCGTTGCCCAGTCCTGCGGCA 184

Qy   197 ATTATCGGGGCATCTTGACCGATTAAATACACACTATCCACCTGCGATATATGCT 256
         ||||
Db   185 ATCAGCGCCGGGGCGGTCACGGCGCCAAGCAGCTACCACCCCCGGCGAGAAGGCGGGCTCG 244

Qy   257 GACAGCTCACTAAAATCTTGACCTTTACCAAGCGCCACCCAAAATCAATGCCAGTGATGAT 316
         ||||
Db   245 CGCAGGTTCATGGAAAATGGGCGGCCCCTTGCCGTCCGCGGGCGGAGCGCAGCCTTGCCG 304

Qy   317 ATGCCATAAAACCGTGCCCAAGCCCATCAAATTGCTCAATGCTTGCAATGGTGGCCGGGGT 376
         ||||
Db   305 TCGATGT---CGGCACCCAGCCCTCCGATCGCGCCCAGGCGCGGGCGGCCGAGGTTGGTGGCC 361

Qy   377 TTTGAATCATTAAAATATGCCTTGCTATCAATATCATCGAATATTCACAGGCCCATG 436
         ||||
Db   362 TTGGAATCGTCGTAGTCACGCGTCAACGCCCTGCCCGTCGGGTACCGAGATGGGAGCC 421

Qy   437 AGCCCTTTGAAAGTTCTTAAGGTATCAAGACATGGCAATCCATCGGGACATTGGCAGCTGTG 496
         ||||
Db   422 AGGCCCGGAAAACGCCTTCAGCCGCCGGAGCAATGGGCGGTCGCAGGCGGACCGGACATGG 481

Qy   497 CCTTAGGCAAGTGCAGATAAGGCAATTGAGTGAGTAGGGTTATGCTTGCCTTTGATTTTTAATTGA 556
         ||||
Db   482 CCCAGCGCCAGCGCCGACGGCGTTGGAATAGTTCACAGGATCGCTTCAGTTCG 541

Qy   557 TGGATTGGTAAAAGTTTAGTTGTTTTTTGATGATATAATGCCATACCATCATCAAAT 616
         ||||
Db   542 CCAACGCGAGCGCAGCTTGTCGAACTGGAACGCCAGCACTTCTGGCCGTCTTCCTCGATC 601

Qy   617 AAATAAAAATCTGCGCGTTTGGTGGCTAAGTAAGCTATTTGTTGTGATTT 676
         ||||
Db   602 AGGCCGAAAGC------------CTTGAAGTCCGGCGGCTTGTTGTTCAGGCCGAA 639

Db   677 GGGGCAACCGCTTGTCGTCGTGTGGGTCAGTCATTGCATCGTCTCTAAAATTGGCATCGTCTCTTGGTAG 736

Db   640 CGACCAGCACGGACGGTATCGGCGTATCGGCGATCAGCGGGGTCGGGTCA--GGGCATCGGCGCGGATTC 697

Qy   737 ATGATGGCGGTTTGGCAATTTTCAAAAAATAGGCAATTTGTGCCAAATAAATTCACCACATG 796
         ||||
Db   698 ACCACGACCTGGCGCGGGAGAGATCCGGTGGATAGTTCACAGGATGCGGGT 757

Qy   797 CCATCGTGCCGATCAAAGATGATCAGGTGAGAGATTTAAAATTGTTTCCCCTTGAGCTGCC 856
         ||||
Db   758 CCGTCGTAGCGGATCCATATGGTCTTCGCTGAGCAGTGCGTCGCCACTGGGTTG 817

Qy   857 AAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCAGCACCACCAAAATCCATATTT 916
         ||||
Db   818 AGGGCGGATCGCAGGTTTCCCAGCTGGAGCTCCAACACGGTACAGCTCGAGTGTCG 877

Qy   917 TGGATCGTTAATAATTCAAGTAGTGCAGGCGGTGCCAATATTACCGGCCCGGACCAACAATCATG 976
         ||||
Db   878 TCG---GCCAGCAGTCGGAGCGCCCGGGGTTGCCGCGGGGACACGCTTG 934

Qy   977 CCTGCATGTTTTTGCCATCTCGCCTACTAAATCGTAACAGTTGCTTTTTGGCATTTGAGCCT 1036
         ||||
Db   935 TCCGGGCGGCCCACCGCCCATTTCGGCCCAGCGGTGGCTCTCTTCGGCGTTGGAACCG 994

Qy  1037 GTGATGGGCCGATGAATTGGGTGTGGTTTG 1061
         ||||
Db   995 GTGATGGGCCGACGATCGGGGGCCTTCG 1019
```

RESULT 10
US-09-701-229-1/c
; Sequence 1, Application US/09701229
; Patent No. 6890910
; GENERAL INFORMATION:

```
Qy    17 TTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATCCATGCTAGCACAAGCA 76

Db    14 TTGCGGAACAGGGCGTTCCGGTTCTTCGAACAGTTCCAGGCTTCGCGGAGGCC 73

Qy    77 GGTGACAAAAGCACAGCCCTAGCGCGTGCGTCCCAATACTCTGCCCGGATCATCAAGT 136
```

; APPLICANT: El-Sherbeini, Mohammed
; APPLICANT: Azzolina, Barbara
; TITLE OF INVENTION: MURD PROTEIN AND GENE OF PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA
; FILE REFERENCE: 20193P
; CURRENT APPLICATION NUMBER: US/09/701,229
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/087,308
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: PCT/US99/11585
; PRIOR FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-701-229-1

Query Match   8.5%;   Score 117;   DB 3;   Length 1450;
Best Local Similarity   48.4%;   Pred. No. 3.6e-25;
Matches 506;   Conservative   0;   Mismatches 500;   Indels 39;   Gaps 5;

```
Qy        17 TTAACAAACTTATCACCTCGCTCATTATAATTATAATTTTTAAATTGATCCATGCTAGCACAAGCA     76
             |   ||||||| |||| ||  | | | ||   |||| |  | ||  |  |  | | ||  ||   ||
Db      1375 TTGGCGAACAGGCGTTCCGCGCTTCTTCGAAGTTCTTGAACATGTCCAGGCTTCGCGCCAGGCC   1316

Qy        77 GGTGACAAAAGCACAGCCCTAGCGCGGTTGCTTTGACTGGCTGCCAATACTGCCCGATCAAGT    136
             |    |  | |  || ||    | | ||   |  || | || |||| || || ||| ||
Db      1315 GGGGACACAGACACCGCATCGCCTCGCC-------GGGCAGCTCGGCGGGCCTGCCCG    1265

Qy       137 GTACCAACCTGATATATTTTAGATGCCAAAATGTGCATCACCTTTGAGTAAATCTTGCTCA    196
             ||| | || |  | | ||| |  || |  |  |||| ||||| ||| |  |  | | |||
Db      1264 ACTGCTTCGTCCAGCGTTCGCGACCGCCACCAGCGCCGTTGCCCTGTTGCCCTGTTGCCCTGGGCA   1205

Qy       197 ATTATCGGGGCATCTTGACCGGATTAAATACACACTACTTCCACCCAGTGAACATATGCT    256
             | ||| ||  |  |  | | |  || | |  |  | ||| ||   | ||  |||  |
Db      1204 ATCAGCCCCGGCGTCACGGCCCAAGCACTACCACCCGTTGCCCAGTGCTTGGGCA   1145

Qy       257 GACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCCACCCAAAATCAAATGCCAGTGATGAT    316
             | |  | |  |  |   ||| |  |  ||  ||   | | |   | |  |  |  | |  ||
Db      1144 CCGCAGGTCATGGAAATCGGCGCCCTTGCGCTCTCCGCGCGAGCAGCACCAGCTTGCCG   1085

Qy       317 ATGCCATAAACCGTGCCAAGCCCATCAATTGCTCGCAAATGGTTGAGCGATGGTTGTACCT    376
             | ||  | | ||   |||||||| ||| | ||| |  || || |  |  |   | || |||
Db      1084 TCGGATGT---CGGCACCCCTCGATCGCCCAGGGCGCGGGCGCCCCGACGTTGGTGGCC   1028

Qy       377 TTTGAATCATTAAAATATGCCTTGCTATCACAATATCAATATATTCAGCGATGCGGT    436
             |||||| | | || | ||   |  || ||| | | | || | |||| |  | |||| ||
Db      1027 TTGGAATCGTCGTAGTAGCTCACCGCCCTGGGCGTACCACTGGCCACGGCGTGAGCC    968

Qy       437 AGCCCTTTGAAAGTCTTAAGGGTATCAAGCAATGGCCATGGGAGAGATTGGCAGCTGTG    496
             | ||| ||| | | ||| |||  | |  |  ||||  |  |||  || |  | |  |  |||
Db       967 AGGCCGGAAAACGCCTTCAGCGCGCCGACGGCATGGGCGTCGAACGCAGCGCCGACCGCATCG    908

Qy       497 CCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCCTTTGATTTTTAATTGA    556
             || ||||| | | |||| | | || | | | |||| || ||| |||| ||| | |  || ||
Db       907 CCCAGCCGCGCGGGCGGCGTTGGAATAGTTGTGGGCGCGTTTCCAGCTGCTGGAATAGCTTCG    848

Qy       557 TGGATTGTAAAAGTTTGTTTTTTGATGATAATGCCATACCATCACCATCAAAT    616
             || |||||||  ||| || |||| ||||| |||| ||||| || ||| ||| ||
Db       847 CCAACCGGCAGCGTGTCGGAACGCAGCCACTTCTGGCCGTCTTCCCTGGATC    788

Qy       617 AAATAAAAATCTGCCGTTGGCTAAGTAATGGTGTGATTT    676
             || | | || |  | |||||  |||| | || | | || |||
Db       787 AGGCCGAAGC------------CTTGAAGTCGGCGCTAAAATTGGCACGTGTCTTCAGGCGTAG    736

Qy       677 GGGGCAACGCCTTGTCAGTGTGGGTCAAGCAATTGGCACGTTCTTGGTAG    736
             ||| | | |  |   | |  | |  |  ||  | | || | | |  |  |||
Db       749 CGACCAGCACGGGCGTATCGCCGATCAGCGGTTCGGTT    692

Qy       737 ATGATGGCGGGTTGGGCAATTTTCAAAAAAATACGCAATTTTTGTGCCAAATAAATCCACCATG    796
             | | |   | ||  |  |   ||   ||| |  | || | |  | | |  |  || | ||  ||||
Db       737 ATGATGGCGGGTTGGGCAATTTTCAAAAAAATACGCAATTTTTGTGCCAAATAAATCCACCATG    796
```

RESULT 11
US-09-489-039A-5247/c
; Sequence 5247, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5247
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5247

Query Match   7.3%;   Score 100.6;   DB 3;   Length 1416;
Best Local Similarity   51.1%;   Pred. No. 3.9e-20;
Matches 337;   Conservative   0;   Mismatches 289;   Indels 33;   Gaps 3;

```
Qy       722 GCATCGTCTTGGTAGATGATGGCGGGTTGGCAATTTCAAAAATACGGCAATTTTTGTGCC    781
             ||  || | || | |  | |  | || ||  |  || |  || | | |  ||||| ||| ||
Db       797 CCATCGTCCGGATCAAGATGATGATCAGCGGAGAGAGATTAAAATTGTGCCCCTTGAGCTGCC    746

Qy       782 AAATAATCACCACCATGCC---ATCGTGCCGATCAAGATGATCAGCGGAGAGATTAAAATT    838
             |||| | |   | ||| |||    | |||||| || || |||| | |  |||| |  | || |
Db       746 GCATCGTCCCGGTTCACCAGCAGTCTTCGGCCATTCTCGTAAATCCGCAGTTTCCGCGCG    687

Qy       839 GTTGCCCTTGAGCTGCCAAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGC    898
             |||||| |||||  ||| |   | | || ||| | |   |  | | |  | |  ||  |||
Db       686 CGATACTGCTGCAGCGCCCAGCGGTAGCCGTTCCATATGGTCGCTTCACGTTGAGGATG    627

Qy       899 ACCACCAAATCCATATTTTGGATCGTTAATAATTCAAGTAGTGCGGCGTGCCAATATTACCG    958
             || ||| | |  ||| | |  | ||| | |  ||| || |  | | | |  |  |  || |
Db       626 GTTGCCGCCACCGCCTGCAGGCTGGTTTCCAGCTGCTGGAAGCTGGAGAGCTCCAGC    567

Qy       959 CCGACACCAACAATCATGCCTGCATGTTTTGCCAGTCTCGCCTACTAAATGTCGTAACAGTG   1018
             || | | | | |  || |  || | || |  | |  | | |   ||  |  | | | ||| |
Db       566 ACATACAGCTTCGACGCGTCCAGCAGCATC---AGCCGCGGCGAGGCGGCGGGATAATTGCCG    510

Qy      1019 CTTTTGGCATTTGACCCTGTGATGCGGATGACGTGATTGGTGTGGTTTGTGCTTGAGGTTAGATCT   1078
             |  | || |||| | | | | || | | | | | |||  | | |  |  |  || |  | | ||
Db       509 CCCACGGCGAGGTTAACCTCCAGCGCTTTCGCCAGCATCTTCGGCCAGGGCGGTTAATGCCG    450

Qy      1079 CGTGCTTGAGTGTATCAATGAACAATTGAACACGCTGATCACAGGGATGCCTTGAGCT   1138
             ||| |  |||| | | |  | | |   || |  ||| | |  | ||   |  || || |
Db       449 CTTTTACCGTTAGAACCGGTAATGGCGGTAATGCGGCGCTTGCGG-------  400
```

```
Db   399   ------------------GCAGACAGTTCGATATCCGCCAACGATCTCTACGCCAGCGTCG   357
Qy  1139   TTGGCGCCGACAATACTCGGCGCGGTTCTTGGGGTTAATGCCGGGGGCTGATGATGATCTGATCG   1198
Db   356   GCCGCGTGGCGTGAGCGAAGGATGGGGCGCAATTCCCGGACTGGCGACAATCAGGTCC   297
Qy  1199   GCTGAGCTTAAGAGTTCACTGTCCAGACGCTGCCAAAATGACAATGACAGCCTGTGGGCAGT   1258
Db   296   GCCGCCAGCAGCCAGGTATCATTGAGTCGCCGACATGGCAATTCGACCGATTCGGGTAAT   237
Qy  1259   TTATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGTCTTAACTTTGTGTGACCGTGATTG   1318
Db   236   TTATCCAGCCCCCGGAGCGCGCCGGTCATGCCCCAGCCCAAATTATGACGACTTTTTTT   177
Qy  1319   ACCAAAAAAATTCACAGCTGACAGTCGCATGCGCAGAGCTTGCCCAAAACCGACGACGGGCATATTTTTTT   1377
Db   176   ATGAAAAGTCCACGCATGACACGGCCCGGTCCATCAGGGCGGGGGTTACCGCCGCC   118
```

```
RESULT 12
US-09-790-988-1
; Sequence 1, Application US/0979098B
; Patent No. 6632935
; GENERAL INFORMATION:
;   APPLICANT: SHIGENOBU, SHUJI
;   APPLICANT: WATANABE, HIDEMI
;   APPLICANT: HATTORI, MASAHIRA
;   APPLICANT: SAKAKI, YOSHIYUKI
;   TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
;   FILE REFERENCE: 08135/0159
;   CURRENT APPLICATION NUMBER: US/09/790,988
;   CURRENT FILING DATE: 2001-02-23
;   PRIOR APPLICATION NUMBER: JP2000-107160
;   PRIOR FILING DATE: 2000-04-07
;   NUMBER OF SEQ ID NOS: 7
;   SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 1
;   LENGTH: 640681
;   TYPE: DNA
;   ORGANISM: Buchnera sp.
US-09-790-988-1

Query Match     6.1%;   Score 84.8;   DB 3;   Length 640681;
Best Local Similarity   50.2%;   Pred. No. 8e-14;
Matches 289;   Conservative 0;   Mismatches 257;   Indels 30;   Gaps 2;
```

```
Qy   802   GTGCCGATCAAGATCATCAGCGGGGAGAGATTTAAAATTGTTGCCCCTTGAGCTGCCAAATT   861
           |||||||||||||||||||||   ||||||||   ||||||||   |||||||   |||||
Db 240727   GTATCGATTGATGATCATCTTCGCTAATATTAAGAATAACGCTACTTAGACTTTTAAATT   240786
Qy   862   TGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCACCACCAAATCCATATTTTTGGAT   921
           |||||||||   ||||||||||   ||||||||   |||||||||||   ||||||   |
Db 240787   AAATGTATTTTCTAGTTGAAAAACTAGACAGTTCTATTATGTATATAAAATCTGCTTCTT---T   240843
Qy   922   CGTTAATAATTCAACTGCAGCGGCGTGCCCAATATTACCGCCCGACACCAACAATCATCCCTGC   981
           |||||||||||||   ||||||   |||||   ||||||||||   |||||||||   |||
Db 240844  ATCAAGTATTTCTAGTACGGGAACAACTATATTACCACCTAGAAAAGCTTTATATCCTGA   240903
Qy   982   ATGTTTTTGCCATCTCGCCTACTACTAATGTCGTAACAGTGCCTTTATATGGCCTGTGAT   1041
           |||||||||||   |||||||||||||   |||||   |||||||   |||||||   |
Db 240904  TTTTTTTGCAATTTTTCAATCATGTAGTAGCTACACTACTTTTTCATTGCCTGTTATCCCCTGTTAT   240963
Qy  1042   GGCGGATGATTGGTGTGTTGTTGTTGCTTGAGTTAGATCTCGTGCTTTGAGTGTATCAATGAA   1101
           |||||||||   |||||||||||   ||||||||||||||||||
Db 240964  TGAAATAATAGGACAGGTCACTTCTCTAG------------AAAA   240996
Qy  1102   CAATTGAACATCGCTGATCACAGAGAATGCCTTGACGTTTGCGGGGACACAATACTCGGGCGT   1161
           ||||||||||||||||   |||||||||||||   |||||   ||||||||   |||||
Db 240997   TAGTTCAATATCACTAATAATCTCAATACCTTAGCAAACGTGCTTTAATTAATAAATAGGTTT   241056
Qy  1162   TCTTGGGTTAATGCCGGGGGCCTGAGCTGATGATGATGAATCTTAAGCTTAAGAGTTCACTGTC   1221
           |||||||||||||   |||||||||||   ||||||   |||||   |||||||   |||
Db 241057  AAATGAAGAAATACCCGGACTTATAAACAATTAAACGCTTAGAATCCACTGATGGGTC   241116
```

```
Qy  1222   CAAGCTGCCAAAATGACAATGACAGCCTGTGGGCAGGTTATCTGCCAAGGTAGGGATTGGGG   1281
           |||||||   |||||||||||||||||||||||||   |||||||||||||
Db 241117  TAAACTCCTAAACTATATATTCAATATATTTGAGGAAATTTAATAAAATTGAAGGATGTTT   241176
Qy  1282   ATTACCATCGATGATGCTAACTTTGTGACCGTGATTGACCAAAAAAATTCACAGCTGACAG   1341
           |||||||||||||||||||   ||||||||   ||||||   |||||||||||||
Db 241177   AGATTCATCAATAATTTAGGTTTATTCCTTTTTTTTTAAAAAAATTAAAATGCAAGTAT   241236
Qy  1342   TCCAGAGCTGCCCAAAACCGACGACGGGCATATTTTTTT   1377
           |||||||||||||||||||||   |||||||||||||
Db 241237  TCCTGTTAAACCCATTCCTAAAATTAATATTTTTT   241272
```

```
RESULT 13
US-09-107-532A-1574/c
; Sequence 1574, Application US/0910753A2A
; Patent No. 6582275
; GENERAL INFORMATION:
;   APPLICANT: Lynn A Doucette-Stamm  and David Bush
;   TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;       ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;   NUMBER OF SEQUENCES: 7310
;   CORRESPONDENCE ADDRESS:
;       ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;       STREET: 100 Beaver Street
;       CITY: Waltham
;       STATE: Massachusetts
;       COUNTRY: USA
;       ZIP: 02354
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE: CD/ROM ISO9660
;       COMPUTER: PC
;       OPERATING SYSTEM: <Unknown>
;       SOFTWARE: ASCII
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/107,532A
;       FILING DATE: 30-Jun-1998
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 60/085,598
;       FILING DATE: 14 May 1998
;       APPLICATION NUMBER: 60/051571
;       FILING DATE: July 2, 1997
;   ATTORNEY/AGENT INFORMATION:
;       NAME: Ariniello, Pamela Deneke
;       REGISTRATION NUMBER: 40,489
;       REFERENCE/DOCKET NUMBER: GTC-012
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (781)893-5007
;       TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1574:
;   SEQUENCE CHARACTERISTICS:
;       LENGTH: 1368 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: double
;       TOPOLOGY: circular
;   MOLECULE TYPE: DNA (genomic)
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;       ORGANISM: Enterococcus faecium
;   FEATURE:
;       NAME/KEY: misc feature
;       LOCATION: (B) LOCATION 1..1368
;       LOCATION: (B) LOCATION 1..1368
;   SEQUENCE DESCRIPTION: SEQ ID NO: 1574:
US-09-107-532A-1574

Query Match     6.0%;   Score 83;   DB 3;   Length 1368;
Best Local Similarity   53.6%;   Pred. No. 9.5e-15;
Matches 173;   Conservative 0;   Mismatches 150;   Indels 0;   Gaps 0;
```

```
Qy   312   ATGATATGCCATAAACCGTGCCAAGCCCATCAAATTGCTGCAAATGGTTGAGCCGATGTTTG   371
           ||| ||| ||| ||| ||| |||   ||| |||   ||| ||| ||| |||
```

Db 1069 ATAACACGACCTTAGACGTTTCAAATCACTTAGAGCCATTTTTAGTCGTCTAAAATATTCG 1010
Qy   372 TACCTTTTGAATCATTAAAAATATGCCTTGGCTCTAATATCATCAATATATTCACAGCGAT 431
Db 1009 TTGCTTTTTGAATCATTATAAAATATTTTCTCGCCTTGGATTTCTCCTACATATTGCGTACGAT 950

RESULT 14
US-09-252-991A-7787
; Sequence 7787, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7787
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7787

Query Match   5.1%; Score 69.8; DB 3; Length 567;
Best Local Similarity   55.0%;
Matches 137; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

Qy   327 CCGTGCCAAGCCCATCAATTGCTCCAATGGTTGAGCCGANGTTTGTTACCTTTTGAATCAT 386
Db   262 CGGCACCCAGCCCCTCGATCGCCGCCCAGGGCGCGGCGCTGGGTTGGCCTTGGAATCGT 321

Qy   387 TAAAATATGCCTTGCTACATCAATAATATTCACAGCGATGCGGTAGCCCTTTGA 446
Db   322 CGTAGTAGTCACCGCCTCGCGTACCCACTGGCGACGCGGAGCCAGGCGGAAA 381

Qy   447 AAGTCTTAAGGGTATCAAGCATGGCATCCATGGGAAGAGATTGGCAGCTGGTGCCTAAGGCAA 506
Db   382 ACGCCTTCAGCGACGCGCGAGCCATGGCCGTCGAACGGCCATGGCCAGCGCGCCA 441

Qy   507 GTGCAGATAAAGGCATTGAGTTAGGGTTATGCTTGCCTTGATTTTTTTAATTGATGGATTGGTA 566
Db   442 CGGCGGGCGAGCGGCGTTGGAATAGTTGTGGGGCGCCACGGATCTTCAGTTGCCAACCGGCA 501

Qy   567 AAAGTTTTGT 575
Db   502 GCAGCTTGT 510

; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 PAGE MILL ROAD
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/221,017B
; FILING DATE: 23-DEC-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1182
; FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP1546
; FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PP2911
; FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU98/01023
; FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Monroy, Gladys H
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-813-5600
; TELEFAX: 650-494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 881:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2402 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...2402
US-09-221-017B-881

Query Match   4.9%; Score 67; DB 3; Length 2402;
Best Local Similarity   46.6%;
Matches 334; Conservative 0; Mismatches 365; Indels 18; Gaps 3;

Qy   355 GGTTGAGCCGANGTGTTTGTTACCTTTTGAATCATTAAAAATATGCCTTGCTATCAATATCATC 414
Db 2165 GGTAGAATTTACGTTCGTAGCTTTCGAGTCATTGATATAATCGACACCCTTTACCGCGAGC 2106

Qy   415 AAATATATTCACAGCGATGCGGTAGCCCTTTGAAAAGTCTTAAGGGTATCAAGCATGGCATC 474
Db 2105 AATCTTTTTCCAGTCGATGCGGTACAATCTTGAAAATCCTGCAAGGGCTTCTCGGAATACAT 2046

Qy   475 CATGGGAAGAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAAGGCATTGAGTAGGTTATG 534
Db 2045 ATTCTTAATATCCATAGCTTTGGGCAGCAATGGCAGTTGCCATAGCATTGTTGCCCTGTTGTG 1986

RESULT 15
US-09-221-017B-881/c
; Sequence 881, Application US/09221017B

```
Qy    535  CTTGCCTTTGATTTTTAAATTGATTGGATTGGTAAAAGTTTGTTTTTTGATGATTATAAATGC  594
           ||||||               ||||               ||||    |||  ||||
Db   1985  CATGCCCGGACAGAGCTAAAAGTTCCTCATCCATAACAAAAAGGTGAATTAAGATTCATGAC 1926

Qy   .595  CATACCATCACCATCATCAAATATAAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATT  654
           |||                                    |||
Db   1925  AACCAATTCGTTCTTTTCTTTCATTGATCCAAGCAGTTGTTATTATCCGTTCGAGCTTCATTGC 1866

Qy    655  TGTTGTTGATAATGGTGTTGATTTGGGGGCAACGGCTTGTCAGTGTGGTCAAGCATTGCTTGGC  714
           ||||                 |||
Db   1865  GAACCGGTAGGAAGAC--------GGGCTACAGGAGGATGTTCGGGCTTACCCATCGGGCTGAT 1815

Qy    715  TAAATTGGGCCATCGTCTTGGGTAGAATGATG------GCGGTTTGGCAATTTTCAAAAATACG  768
           |||                        ||||||
Db   1814  AAAAGGATCATCCTCCCAGTAGATGAAGCAATCCTCGGGTTGCTGATTCCGAGTGATTCT 1755

Qy    769  CAATTT----TTGTGCCAAATAATCCACCATCGTGCCGGATCAAGATGATCAGCGGGA  825
           |||           ||                 |||        |||||
Db   1754  CATTTTCGGCTTCTCGCATAAAGCTCGAAATCGGTGATCGTATCGGATCTAAAATGATCCGGTGT 1695

Qy    826  GAGATTTAAAATTGTTGCCCCTTGAGCGTGCCAAAATTTGAGATATGCTCAAGCTGAAAACT  885
           ||||        |||||                 ||||        ||||||||
Db   1694  AATAATTGAGCAGAATGGCCACATTAGCCCTGAAAGTCGTACATGTTGTCCAGTTGAAAAACT 1635

Qy    886  TGACAGCTCAAGCACCACCAAAATCATATATTTGGATCGTTAATAATTCAAGTGCAGGCGT  945
           |||        ||                                  |||||
Db   1634  GCTCAACTCTTATTACATAATAAGGATCGAGGATCGTAAGCGACCTGTCTGGCCAGACTAAA 1575

Qy    946  GCCAATATTACCGCCGACACACAACGAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAA 1005
           |||                        |||||              |||||
Db   1574  ACCGACATTGCCTGCCAAACCTACATCCCAAACCGGCCTTACGAGGGTGTGATACAACCA 1515

Qy   1006  TGTCGTAACAGTGCTTTTGGCATTTGAGCCTGTGATGGCGGATGATTGGTGTGTTTTGT 1062
           |||        |||                 ||||                 |||||||
Db   1514  CATGGTTGTTGTTTGTGTCTTGCCATTACTACCCGTGATGCACACCAATAAATGCATCGGT 1458
```

OM nucleic - nucleic search, using sw model

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: May 12, 2006, 13:01:47 ; Search time 1245 Seconds
(without alignments)
9166.058 Million cell updates/sec

Title: US-10-672-787-35_COPY_11357_12736
Perfect score: 1380
Sequence: 1 taaagcgtgaaccatgttaa............acgacgcatattttttga 1380

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 9793542 seqs, 413689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications NA Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1380	100.0	1398	7	US-10-282-122A-27195	Sequence 27195, A
2	1380	100.0	96109	7	US-1-672-787-35	Sequence 35, Appl
3	231	16.7	1344	7	US-10-282-122A-8940	Sequence 8940, Ap
4	160.4	11.6	1314	3	US-09-815-242-7090	Sequence 7090, Ap
5	160.4	11.6	1314	9	US-10-282-122A-22236	Sequence 22236, A
6	160.4	11.6	11498	3	US-10-958-216-400	Sequence 400, App
7	160.4	11.6	1314	9	US-09-754-468-44	Sequence 44, Appl
8	160.4	11.6	1830121	7	US-10-329-670-1	Sequence 1, Appli
9	160.4	11.6	1830121	9	US-10-158-865-1	Sequence 1, Appli
10	160.4	11.6	1830121	9	US-10-981-687-1	Sequence 1, Appli
11	154.4	11.2	14324	7	US-10-795-159-569	Sequence 569, App
12	152.8	11.1	908766	9	US-10-795-159-685	Sequence 685, App
13	152.4	11.1	1281	7	US-10-282-122A-25165	Sequence 25165, A
14	152.4	11.0	1314	9	US-10-958-216-402	Sequence 402, App
15	142.2	10.3	1305	7	US-10-282-122A-30713	Sequence 30713, A
16	139.6	10.1	1317	3	US-09-741-669-259	Sequence 259, App
17	139.6	10.1	1317	9	US-09-815-242-5932	Sequence 5932, Ap
18	139.6	10.1	1317	7	US-10-282-122A-20229	Sequence 20229, A
19	139.6	10.1	28277	3	US-09-754-468-2	Sequence 2, Appli
20	123.6	9.0	1308	7	US-10-282-122A-32629	Sequence 32629, A
21	117.8	8.5	1182	9	US-10-282-122A-33528	Sequence 33528, A
22	117	8.5	1347	3	US-09-815-242-7912	Sequence 7912, Ap
23	117	8.5	1347	7	US-10-282-122A-30468	Sequence 30468, A
24	113	8.2	1404	7	US-10-282-122A-41274	Sequence 41274, A
25	111	8.0	1317	7	US-10-282-122A-39003	Sequence 39003, A
26	110.8	8.0	4942	9	US-10-450-763-25620	Sequence 25620, A
27	110.8	8.0	4944	9	US-10-450-763-25077	Sequence 25077, A
28	110.8	8.0	4944	9	US-10-450-763-25985	Sequence 25985, A
29	110	8.0	1350	7	US-10-282-122A-31716	Sequence 31716, A
30	109.4	7.9	1317	3	US-09-815-242-9978	Sequence 9978, Ap
31	109.4	7.9	1317	9	US-10-282-122A-39835	Sequence 39835, A
32	103.8	7.5	1314	7	US-10-282-122A-23400	Sequence 23400, A
33	94.2	6.8	1314	7	US-10-282-122A-19656	Sequence 19656, A
34	92.8	6.7	2058	9	US-10-450-763-8816	Sequence 8816, Ap
35	92.8	6.7	4831	9	US-10-450-763-30346	Sequence 30346, Ap
36	92.6	6.7	1317	9	US-10-450-763-41740	Sequence 41740, A
37	89.6	6.5	3117	9	US-10-450-763-4899	Sequence 4899, Ap
38	89.6	6.5	3117	9	US-10-450-763-25044	Sequence 25044, A
39	88.6	6.4	1353	9	US-10-282-122A-17455	Sequence 17455, A
40	84.8	6.1	640681	3	US-09-790-988-1	Sequence 1, Appl
41	83	6.0	1359	7	US-10-282-122A-21621	Sequence 21621, A
42	81	5.9	1071	7	US-10-282-122A-37050	Sequence 37050, A
43	73	5.3	1338	7	US-10-282-122A-29879	Sequence 29879, A
44	70.2	5.1	1440	7	US-10-282-122A-17840	Sequence 17840, A
45	67	4.9	2402	5	US-10-194-163-881	Sequence 881, App

ALIGNMENTS

RESULT 1
US-10-282-122A-27195/c
Sequence 27195, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,338
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patentin version 3.1
SEQ ID NO 27195
LENGTH: 1398

```
;  TYPE: DNA
;  ORGANISM: Moraxella catarrhalis
US-10-282-122A-27195

Query Match      100.0%;   Score 1380;  DB 7;  Length 1398;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1380;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TAAAGCGTGAACCATGGTTAACAAAACTTATCACCTCGCTCATTATAATAATTTTAAATTGATC 60
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1383 TAAAGCGTGAACCATGGTTAACAAAACTTATCACCTCGCTCATTATAATAATTTTAAATTGATC 1324

Qy   61 CATGCTAGCACAAGCAGGTGACAAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 120
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1323 CATGCTAGCACAAGCAGGTGACAAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 1264

Qy  121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAATGTGCATCATCCTTT 180
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1263 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAATGTGCATCATCCTTT 1204

Qy  181 GAGTAAATCTTGCTCAATTATTCGGGGCATCTTGACGGATTAAATACACTATTCCACCCA 240
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1203 GAGTAAATCTTGCTCAATTATTCGGGGCATCTTGACGGATTAAATACACTATTCCACCCA 1144

Qy  241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCCAAAAT 300
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1143 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCCAAAAT 1084

Qy  301 CAATGCCAGTGATGATGATTGCCATAAAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1083 CAATGCCAGTGATGATGATTGCCATAAAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 1024

Qy  361 GCCGGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1023 GCCGGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 964

Qy  421 TTCACAGCGGATGCGGTAGCCCTTTGAAAAGTTCTTAAGGGTATCAAGCACATGGCATCCATGGG 480
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  963 TTCACAGCGGATGCGGTAGCCCTTTGAAAAGTTCTTAAGGGTATCAAGCACATGGCATCCATGGG 904

Qy  481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGTGCTTTGCC 540
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  903 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGTGCTTTGCC 844

Qy  541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATAATGCCATACC 600
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  843 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATAATGCCATACC 784

Qy  601 ATCACCATCATCAAATAAATAAAATAAAATCTCGCCGTTTGGGCTAAGTAAGCTATTGTTGT 660
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  783 ATCACCATCATCAAATAAATAAAATAAAATCTCGCCGTTTGGGCTAAGTAAGCTATTGTTGT 724

Qy  661 GATAATGGTTGTGATTGGGGCAACGCTTGTCAGTGTGGTGGTCAAGCGTTCAGTTCAAGATT 720
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  723 GATAATGGTTGTGATTGGGGCAACGCTTGTCAGTGTGGTGGTCAAGCGTTCAGTTCAAGATT 664

Qy  721 GGCATCGTCTCTTGGTAGATGATGGCGGTTGGCAATTTCAAAAAATACGCAATTTTTGTGTC 780
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  663 GGCATCGTCTCTTGGTAGATGATGGCGGTTGGCAATTTCAAAAAATACGCAATTTTTGTGTC 604

Qy  781 CAAATAAATCCAACATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTGT 840
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  603 CAAATAAATCCAACATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTGT 544

Qy  841 TGCCCCTTGAGCTCCGCCAAAATTTGAGATGATGCCTCAAAAACTTGACAGCTCAAGCAC 900
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  543 TGCCCCTTGAGCTCCGCCAAAATTTGAGATGATGCCTCAAAAACTTGACAGCTCAAGCAC 484

Qy  901 CACCAAAAATCCATATTTTGGATCGTTAAATAATTCAAGTGCAGCGCAGGCGGTGCCAATATTATTACCGCC 960
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  483 CACCAAAAATCCATATTTTGGATCGTTAAATAATTCAAGTGCAGCGCAGGCGGTGCCAATATTATTACCGCC 424

Qy  961 GACACCAACAAATCATGCTGCATGTTTCGCCATCTCGCCTACTAATGTCGTAACAGTGCT 1020
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  423 GACACCAACAAATCATGCTGCATGTTTCGCCATCTCGCCTACTAATGTCGTAACAGTGCT 364
```

```
Query Match      100.0%;   Score 1380;  DB 7;  Length 96109;
Best Local Similarity   100.0%;   Pred. No. 0;
Matches 1380;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 TAAAGCGTGAACCATGGTTAACAAAACTTATCACCTCGCTCATTATAATAATTTTAAATTGATC 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11357 TAAAGCGTGAACCATGGTTAACAAAACTTATCACCTCGCTCATTATAATAATTTTAAATTGATC 11416

Qy    61 CATGCTAGCACAAGCAGGTGACAAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11417 CATGCTAGCACAAGCAGGTGACAAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 11476

Qy   121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAATGTGCATCATCCTTT 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11477 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAATGTGCATCATCCTTT 11536

Qy   181 GAGTAAATCTTGCTCAATTATTCGGGGCATCTTGACGGATTAAATACACTATTCCACCCA 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11537 GAGTAAATCTTGCTCAATTATTCGGGGCATCTTGACGGATTAAATACACTATTCCACCCA 11596

Qy   241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCCAAAAT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 11597 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCCAAAAT 11656
```

Right-column alignment continuation (Db 96109 reference):

```
Db  423 GACACCAACAAATCATGCTGCATGTTTCGCCATCTCGCCTACTAATGTCGTAACAGTGCT 364
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Qy 1021 TTTGGCATTTGAGCCTTGTGATGGCCATGATGGTGTGTTTGCCATCCTGCCTACTAATGTCGTCGTAACAGTGCT 1080
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  363 TTTGGCATTTGAGCCTTGTGATGGCCATGATGGTGTGTTTGCCATCCTGCCTACTAATGTCGTCGTAACAGTGCT 304

Qy 1081 TGCTTTGAGTGTATCAAATGAACAATGAACATGCGCTGATCACGCTGTTGTTGCTTGAGTCGGC 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  303 TGCTTTGAGTGTATCAAATGAACAATGAACATGCGCTGATCACGCTGTTGTTGCTTGAGTCGGC 244

Qy 1141 GGCGGCGGACAATACTCGGGCGTTCTTGGGGTTAAATGCCGGGCGCTGATGATGGTTT 1200
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  243 GGCGGCGGACAATACTCGGGCGTTCTTGGGGTTAAATGCCGGGCGCTGATGATGGTTT 184

Qy 1201 TGAGCTTTAAGAGTTCACTGTCCAACCTGCCAAAAATGACAATGAACGCCCCTGGGCAGTTT 1260
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  183 TGAGCTTTAAGAGTTCACTGTCCAACCTGCCAAAAATGACAATGAACGCCCCTGGGCAGTTT 124

Qy 1261 ATCTCGCCAAGGTAGGATTGGGATTACCCATCGATGATGATGCTAACTTGTTGTGAGTTGAC 1320
        |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  123 ATCTCGCCAAGGTAGGATTGGGATTACCCATCGATGATGATGCTAACTTGTTGTGAGTTGAC 64

Qy 1321 CAAAAAAATTCACAGCTGACAGTCCAGAGCGTGCCCAAAACCGACGAGACGGGCATATTTTTTGA 1380
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   63 CAAAAAAATTCACAGCTGACAGTCCAGAGCGTGCCCAAAACCGACGAGACGGGCATATTTTTTGA 4
```

```
Qy    301 CAATGCCAGTGATGATGATGCCATAAACCGTTGCCCAGCCCATCAATTGCTCGAAATGGTTGA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11657 CAATGCCAGTGATGATGATGCCATAAACCGTTGCCCAGCCCATCAATTGCTCGAAATGGTTGA 11716

Qy    361 GCCGGATGTTTGTTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11717 GCCGGATGTTTGTTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 11776

Qy    421 TTCACAGCGATGCGGTAGCGCCCTTTGAAAGTCTTAAGGGTATCAAGCATCCATGGG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11777 TTCACAGCGATGCGGTAGCGCCCTTTGAAAGTCTTAAGGGTATCAAGCATCCATGGG 11836

Qy    481 AAGATTGGCAGCTGTTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11837 AAGATTGGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGGCATTGAGTAGGTTATGCTTGCC 11896

Qy    541 TTTGATTTTTAATTGATTGGATTGGGTAAAAGTTTGTTTTTTTGATGATAATGAATGCCATACC 600
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11897 TTTGATTTTTAATTGGATTGGGTAAAAGTTTGTTTTTTTGATGATAATATAATGCCATACC 11956

Qy    601 ATCACCATCATCAAATAAAATAAAAATCTGCCGTTTTGGTTTGGTGGCTAAGTAAGCTATTTGTTTGT 660
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11957 ATCACCATCATCAAATAAAAATAAAAATCTGCCGTTTTGGTGGCTAAGTAAGCTATTTGTTTGT 12016

Qy    661 GATAATGGTGTGATTTGGGGGCAACGCTTGTCGTCAGTGTGGTCAAGCATTGCTTGGGCTAAATT 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12017 GATAATGGGTGTGATTTGGGGGCAACGCTTGTCGTCAGTGTGGTCAAGCATTGCTTGGGCTAAATT 12076

Qy    721 GGCATCGTCTTGGTAGATGAGGGGGTTGGCAATTTCAAAAATACGCAATTTTTGTGC 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12077 GGCATCGTCTTGGTAGATGAGGGGGTTTGGCAATTTCAAAAATACGCAATTTTTGTGC 12136

Qy    781 CAAATAAATCCACCATGCCCATCGTGCCGATCAAGAGATCAGCGGAGAGATTTAAAAATTGT 840
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12137 CAAATAAATCCACCATGCCCATCGTGCCGATCAAGATGATCACGCGGAGAGATTTAAAAATTGT 12196

Qy    841 TGCCCCTTGAGCTGCCAAAATTTGAGATAATGCTCAAGCTGAAAACTTGACAGCTCAAGCAC 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12197 TGCCCCTTGAGCTGCCAAAATTTGAGATAATGCTCAAGCTGAAAACTTGACAGCTCAAGCAC 12256

Qy    901 CACCAAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGCGGTGCCAATATTACCGCCC 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12257 CACCAAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGCGGTGCCAATATTACCGCC 12316

Qy    961 GACACCAACAATCATGCCTGCATTGTTTCGCCATCCTCGCCTACTAATGTCGTPAACAGTGCT 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12317 GACACCAACAATCATGCCTGCATTGTTTCGCCATCTCGCCTACTAAGTGTCGTPAACAGTGCT 12376

Qy   1021 TTTGGCCATTTGAGCCTGTGATGGGCGATGATTGGTGTGTGTTTTGTTGCTTGAGTTAGATCTCG 1080
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12377 TTTGGCCATTTGAGCCTGTGATGGGCGATGATTGGTGTGTGTTTTGTTGCTTGAGTTAGATCTCG 12436

Qy   1081 TGCTTGAGTGTATCAATGAACAATTGAACATGCGTCGATCACAGGAATGCCTTGAGCTTT 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12437 TGCTTGAGTGTATCAATGAACAATTGAACATCGCGTGATCACAGGAATGCCTTGAGCTTT 12496

Qy   1141 GGCGGCGACAATACTCGGCGCTTCTGGGTTAATGCCGGGGCTGGGGGGGCTAAATCGGC 1200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12497 GGCGGCGACAATACTCGGCGCTTCTGGGTTAATGCCGGGGCTGGGGGGGCTGATCAATCGGC 12556

Qy   1201 TGAGCTTAAGAGTTCACTGTCTCCAAGCTGCCAAAATGACAATGAACGGCCTGTGGGGCAGTTT 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12557 TGAGCTTAAGAGTTCACTGTCTCCAAGCTGCCAAAATGACAATGAACGGCCTGTGGGGCAGTTT 12616

Qy   1261 ATCTGCCAAGGTAGGATTGGGATTACCATGCGATGATGATGAGGCTTCTTTAGGCGTTGCTGCGCCT 1320
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db  12617 ATCTGCCAAGGTAGGATTGGGATTACCATGCGATGATGAGGCTTCTTTAGGCGTTGCTGCGCCT 12676

Qy   1321 CAAAAAAATTCACAGCTGACAGTCGCCCAAACCCAACCGACGAGGGCACATATTTTTTTTGA 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12677 CAAAAAAATTCACAGCTGACAGTCGCCCAAACCGACGAGGGCACATATTTTTTTGA 12736
```

```
Query Match             16.7%;  Score 231;  DB 7;  Length 1344;
Best Local Similarity   52.4%;  Pred. No. 2.7e-53;
Matches  721;  Conservative  0;  Mismatches 590;  Indels 66;  Gaps  7;

Qy      2 AAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATATTTTAAATTGATCC 61
          ||| || ||| ||  |  |||   | | ||  |  || | ||||  ||||||||| ||
Db   1340 AACGAATTGACGCAGCGAACAAAACTGCTGACCACCGGTCATTATAACTTTTAAACATATCA 1281

Qy     62 ATGCTAGCACAAAGCAGGTGACAAAAAGCACAGCCCCTAGCGGTGTGCTTTGACTGGCTGCCAAT 121
          |||   |   | |||  ||| ||  |||  |    || || | ||| | |  ||  | |||
Db   1340 AACTTGCACATGCTGGTGAT---------AGCAATACCACACATCTTCAGCGTTGTTTCAC 1229

Qy    122 ACTGCCCGATCAAAGTGTACCAAACCTGATATATATTTTAGAATGCGCAAAATGTGCATCACCTTTG 181
          ||   |   |  |    ||  ||  |  | | || || | |  ||| ||  | | || |
Db   1228 GTTGACACAGCTCGACAGCTTCTTTAAGCGTTGCTGCGTATAAATACACAAAATTTTAGTTGCGCCT 1170

Qy    182 AGTAAAATCTTGCTCAATTATCGGGGGCATCTTCTTTAAGCCGTTGACCGATTAAATACACACTATCACCCAC 241
          || || |  || ||   ||   ||||  | |    |   ||
Db   1169 TGAATGGCTTGCTCGATGACCGGCGCCGCCGAATCTTCACCAAATCAATACACACACTATCACCCAC 1110

Qy    242 TGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCACCAAAATC 301
          ||   ||   | | || | | | |||  | ||  |  || |  || ||| ||||| |||
Db    242 TGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCACCAAAATC 301
```

RESULT 4
US-09-815-242-7090/c
Sequence 7090, Application US/09815242
Patent No. US2002061559A1
GENERAL INFORMATION:
  APPLICANT: Haselbeck, Robert
  APPLICANT: Ohlsen, Karl L.
  APPLICANT: Zyskind, Judith W.
  APPLICANT: Wall, Daniel
  APPLICANT: Trawick, John D.
  APPLICANT: Carr, Grant J.
  APPLICANT: Yamamoto, Robert T.
  APPLICANT: Xu, H. Howard
  TITLE OF INVENTION: Identification of Essential Genes in
  TITLE OF INVENTION: Prokaryotes
  FILE REFERENCE: ELITRA.011A
  CURRENT APPLICATION NUMBER: US/09/815,242
  CURRENT FILING DATE: 2001-03-21
  PRIOR APPLICATION NUMBER: 60/191,078
  PRIOR FILING DATE: 2000-03-21
  PRIOR APPLICATION NUMBER: 60/206,848
  PRIOR FILING DATE: 2000-05-23
  PRIOR APPLICATION NUMBER: 60/207,727
  PRIOR FILING DATE: 2000-05-26
  PRIOR APPLICATION NUMBER: 60/242,578
  PRIOR FILING DATE: 2000-10-23
  PRIOR APPLICATION NUMBER: 60/253,625
  PRIOR FILING DATE: 2000-11-27
  PRIOR APPLICATION NUMBER: 60/257,931
  PRIOR FILING DATE: 2000-12-22
  PRIOR APPLICATION NUMBER: 60/269,308
  PRIOR FILING DATE: 2001-02-16
  NUMBER OF SEQ ID NOS: 14110
  SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 7090
  LENGTH: 1314
  TYPE: DNA
  ORGANISM: Haemophilus influenzae
  FEATURE:
  NAME/KEY: CDS
  LOCATION: (1)...(1314)
US-09-815-242-7090

Query Match   11.6%;   Score 160.4; DB 3; Length 1314;
Best Local Similarity   49.6%;   Pred. No. 1.4e-33;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

Qy   1 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATATAATTTTTAAATTGATC 60
              |||| ||||||| ||||||||||||| ||||||||||||||||| |||||| ||||||||
Db   1308 TAAACACATTGAGCTAAATGCGTAAAATCTTCGCCCGCGGCCTTTTCAAAGAAGCAAAACTGAATC 1249

Qy   61 CATGCTTAGCACAAGCAGAGGTGACAAAAGCACAGCCCTAGCGGGTGCTTTGACTGGCTGCCAA 120
              |||| ||||||||||||| |||||||||||| |||||| |||||||||||||||
Db   1248 GAGACTTGCACAAGCAGGCGACAAΤΑΑΤΑCCA------TATCTCCGCTTTGCAA 1201

Qy   121 TACTGCCCGATCAAGTGTACCAAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTTT 180
              |||||||| |||||||||||||||||| ||||| ||||||||||||||||||||||
Db   1200 TGTTGGGCGGTAAAAATTCTATCGCTTGTTCCATTGTATCGGAACAA------ΑΤΑACΤΤΤG 1147

Qy   181 GAGTAAATCTTGCTCAATTATCGGGGGCCATCTTGACCGGATTAAAΑΤΑCACΑΤΑΤCΑCCCA 240
              ||||||||||||| |||||||||||||||||| |||||||| |||| ||||||||||||||
Db   1146 CGATGAAAΑΤΤΤGCAAGCAGCGGCACCΑΤCΤCGACCAAAACAΑΤΑΑCAΑΑΑΤΑΑGΤGΤGG 1087

Qy   241 CTGCATAACΑΤΑΤGCTGACAGCTCACTAAAΑΤCΤΤGACCΤΤΑCCAΑΑΤΑΑCAΑΑΑΤCΑGCCCΤ 300
              |||||||| |||||||||||||||| ||||||||||||||||||| |||||
Db   1086 TTGATTAAΤΤΑAΤΤCAGCΤΑΑΤΤΤCΤGAAAAATCΑGCCCCΤΤΤΤCCGΤCΤCC 1027

Qy   301 CAATGCCAGTGATGATΑΤGCCΑΤΑAAACCGΤGCCAΑGCCΑΤCΑΑΤΤGCΤCΑΑΤGGΤΤGA 360
              ||| |||||||||||||||||||||||||||| ||||||||||||||||||
Db   1026 CAAAΤGCAΑΤΤΤΑCCCΤCAΑΤΑΤΑ------AAGCCCΑGCCAΑΤGCAGCCAACΤGΤΑCΤ 976

Qy   1316 TTGACCAAAAAΑΤΤCACAΓCΓGACAGΤΓCCΑGAGCΓGΓCCCAAΑCCGACGACGGGCAΓΑΤ 1372
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   80 TCAΤGCAGGAAΑΤΤΤΑCAGCAGAGAAACACCΓGAΤΑΤΤCCCAAAGCCΓGCΓΑCAAACΤΤΤ 24

Qy   140 AΤCΤGAΤCGΤGΤΓCGGGAGGΤGΓGGGGΑGΓCCGΓΤΑCΑGACGACGAACΓΓGGΓΑGCCΓΤGA 81
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||

Qy   1256 AGΤΤΤΑΤCΤGCCAAGGΤΑGGAΤΓGGGAΓΤΑCCAΤCGAΤGΑΓGΤCΤΑACΓΤΤGΓΓGACCGΓGA 1315
              |||||||||||||||| ||||||||||||||||||||||||||||||||||||||

Qy   200 ΤCΤCCΓΓGΤΤΑΑΤΑΑΤCΓΓΓGAΓCAΑGΓGΓGAACCΓGΓΑAAAΓGCCΓGΓGΓGGGC 141

Qy   1196 ΤCCGGCΓGAGCΓΓΑAGAGΓΓCACΓCΓGCCAAAΑΤGACAAΓGCGCΓCΓGΓGGGC 1255

Qy   260 GCAAΤAGCΓCGCCΓGAΑΑΓΓΓCCGGΓΑAΓΓGΓGΓGCAAGGCCGCΓΑAAAΓAAΓCΓCΓ 201

Qy   1136 GCΓΓΓGGCGCGGACAAΑΤΑCΓCGGCGCΓΓCΓΓGGGGΓΓΑAΓGCCGCΓGAΓGAΓCΑΓCΓGA 1195

Qy   303 ------------ΑGACGCCAAΤΑAΓΓGGAΓΑΓΓGCΑCCΓΑCΑGAAAΓΑCCΓΓΓΑ 261

Qy   1076 ΤCΓCGΓGCΓΓΓGAGΓGΓΑΓCAΑΓGAACAAΓΓGΑCΑΓCAΓΓGGAAΓGCCΓΓGA 1135

Qy   353 GΓΑCΓΓΓΓΓΓGCAΓΓGΓGAΑGCΓGΓΑAΓCGGCACAΓCAGΓGGC-------- 304

Qy   1016 GΓGCΓΓΓΓΓGGCAΓΓGΑGCGCΓGGAΓGCGGAΓGAΓΓGΓGΓGΓΓCΓΓGAGΓΑGA 1075

Qy   413 CCGGCCΓΑCΓGCAACΓΓΓCΓΓΑCCΓGCAΓCCΓΓΑGCCΑΓΓGAAGΓΓCGGΓΓΑCA 354

Qy   956 CCGCGACCAACAΑΓCAΓGCCCΓGCGCAΓGΓΓΓΓGCCΓCΓΑΓΓGΓCGΓΑACA 1015

Qy   470 AGGACΓΑGΓΓAGCΓC----ΓGGΓΓGAΓCΓΓΓAAGΓΑΑΑΓCΓΓΑACGCΓGΓCGACCAAGAΓG 414

Qy   896 AGCACCACCAΑΑΑΓCCAΓΑΓΓΓGGAΓΓCGΓΓΑΑΓΓCAAGΓGCAΓGCCGΓGCCAΑΓΑΓΓΑ 955

Qy   836 AΓΓGΓΓΓGCCCCΓΓGAGCGCGCCAAΑΓΓΓGAGAΓΑΓGCCGΓΑΓCAGCGAGAGAΓΓΓΑA 895

Qy   530 ACCACΓGCΓΑCCΓCAGCGΓΓΓΓΑAGGΓGGΓΓΓΓCΓΓΑACΓGACAΓCΑGΓΓΑACΓCA 471

Qy   590 GCΓΓGGΓGAΓΑCCCCΑGCAΓΑΓΓCGGΓΓCΓCAΑΑΓΓΓCΑAAΑΑΑΓΑCGCAAΑΓΓΓΓ 531

Qy   776 ΓGΓGΓCCAΑΑΓΑAΓCCACCAΓCGΓGCGAΓCAΑGΑΓGAΓCΑGCΓGACAGΓΓΓΑA 835

Qy   646 ----AAGCAΓCGΓCΓCGGΓΓΑAΑΑΓΑCΑACΓΓΓΓΓCAAΑAΑΓΓΑCΓCΑΓΑΓΓGΑΑ 591

Qy   716 AAAΓΓGGCAΓCGΓCΓGGGΓΑGACGGΓΓΓCCAAΓΓΓCΓGGAACAAGΓGGGGCΓΓΑ- 775

Qy   697 ΓΓAAACCAΑΑGCΓΓΓGCAΓΓGGΓGΓGΓCGΓAΓCΓGACAACAAGΓGGCΓΓΑ----- 647

Qy   656 GΓΓGΓGAΓΑAΓGGΓGΓGΓGAΓΓΓGGGGCCACGCΓΓGΓCAGΓGGΓCAAGCAΓΓGCΓΓGGCΓ 715

Qy   749 AGCGΓΑCCGΓGCGGGCAΓCΓCΓΓAAAACGCCAΓAΓGAΓΓ----ΓΑΑΑΓCΓGGΓGCGΓ 698

Qy   596 AΓΑCCAΓCACCAΓCAΓCAAΑΓAAΑΓACΓCCGΓΓΓGGGCΓGΓAGCΓΑAGCΓΑΓΓΓ 655

Qy   416 AΓΑΓΑΓΓCACAGCGAΓGCGGΓΑGCCCΓΓΓGAΑAGΓCΓΓΑAGGΓΑΓCΑAGCAΓGGCAΓCC 475

Qy   929 ACGΓΑCΓCACAGCGGAΓGCΓCΓΑAΓCCΓΓΓAAAΓCGΓΓΓΓΑAAGΓΓΓCAAGCAΓGACΓCC 870

Qy   476 AΓGGGAAGAΓΓGGCAGCΓGΓGCCΓΑAGGCAAGΓGCAGAΓΑAGGGCAΓΓGAGΓΑGGΓΓAΓGC 535

Qy   869 AΓΑGGΓΑACCAAΓΓGCCΓCCCCAAΑΓGCΓΑAACACGCΓΓAAAGCAΓΓGGCΓΑCGΓΓGΓGC 810

Qy   536 ΓΓGGCΓΓΓGAΓΓΓΓΓΓΑAΓΓGAΓΓGGAΓΓGGΓΑΑAAGΓΓΓGΓΓΓΓΓΓGAΓGAΓΑAΑΓGCC 595

Qy   809 AΓΑCCΓΓGAΑΓΑΓΑΓΑAAΑΓCCGAGCΓΓΓΓΑAΓΓΑACCGCΓGΓΑAACCACGAGCAAGCCAA 750

Qy   356 GΓΓGAGCCGAΓGΓΓΓGΓΑCCΓΓΓΓGAAΓCAΓΓAAΓCAΓΓΑAAΓΑΓGCCΓΓGCΓΑΓCAAΓΑΓCA 415

Qy   989 GΓΑGCACCAΑCAΓΓGGΓΑCCΓΓΓAGAGΓCAΓΓΑΓΑΑΓΑACGGGACΑΓCAGCACCAGΓΓΓΑ 930

Qy   302 AAΓGCCA------GΓGAΓGAΓΑΓGGCCΑΓΑAAACCGΓGCCAAGCCCΑΓCAAΓΓGCΓCAAΓG 355

Qy   1049 AGΓGCAACCΓΓΑCCΓΓΓΓCACΓΓCAAΓGCAGCACCΓΓΑGCCΑΓΓΑACΓΓCΑΓGCΓGCAAGΓ 990

Qy   1109 ΓΓΓΓCGAΓΑGAΑGAACGΓΑAAΑGGAGAΑAAAAΓCΓΓGCCCΓΓΓGCCCΓGΓCCACCΓΑAAΑΓΑ 1050

```
; RESULT 5
; US-10-282-122A-22236/c
; Sequence 22236, Application US/10282122A
; Publication No. US20040029129A1
```

```
  Query Match             11.6%;  Score 160.4;  DB 7;  Length 1314;
  Best Local Similarity   49.6%;  Pred. No. 1.4e-33;
  Matches  683;  Conservative   0;  Mismatches  606;  Indels  89;  Gaps  7;

Qy      1 TAAAGCTGAACCATGTTAACAAACTTATCACCTCCGCTCATTATATAATTTAAATTGATC 60
            ||||    |||||  |||  |  || |||| |  | ||| | | |  | || || ||||
Db     1308 TAAACATTGAGCTAAATGCGTAAAATTCTTCGCCGCGCCTTTTCAAAGAGAGAAACTGATC 1249

Qy     61 CATGCTAGCACCAAGCAGAGGTGACAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 120
            |||| |||   |||  || ||| ||||  |||  ||||   |  |  |||| ||  ||||
Db    1248 GAGACTTGCACAGACGAGGCGACAATAATACCA------TATCTCGCTTTGCAA 1201

Qy    121 TACTGCCCGATCAAGTGTACCAAACCTGATATATTTTTTAGATGATCCAAANTGCATCACCTTT 180
            |||| |||| |||| ||| ||| |||||| | | ||||| | ||| | |  |||||| |||
Db    1200 TGTTGGGCGTAAAAATTCTATACGCTTGTTCCCATTGTTATCGAAACAA------ATAACTTTG 1147

Qy    181 GAGTAAATCTTGCTCAATTATCGGGGCATCTTGTTACCGATTAAATACACACATATCACCCA 240
            ||||  |    |    |||   || | |||||| ||| |||| | ||||||   ||||| |
Db    1146 CGATGAAAATTTGCAAGCAGCGCACCATCGTCGACCAAAACAAATAACAAAATAAATGTGTGG 1087

Qy    241 CTGCATAACATATGCTGAGACGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
            |||  |  ||||| | |  ||||||| ||| |    |||  || ||  |||| | |||
Db    1086 TTGATTAATTAATTCAGCTAATTCTGAAAAAATCAGCCCCTTTTCCGTCTCCTCCGCTTAGCAA 1027

Qy    301 CAATGCCAGTGATGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAAATGGTTGA 360
            |||   |  |||  |  ||||| | | || |   | |   ||||| | | | |  |||| ||||
Db    1026 CAAAATGCAATTTACCCTCAATATA-------AAGCCCAGCCAATGCAGCAACTGTACT 976
```

RESULT 6
US-10-958-216-400/c
; Sequence 400, Application US/10958216

Qy   Query Match   11.6%;   Score 160.4;   DB 9;   Length 1314;
      Best Local Similarity   49.6%;   Pred. No. 1.4e-33;
      Matches 683;   Conservative 0;   Mismatches 606;   Indels 89;   Gaps 7;

Qy   1 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATAAATTTTTAAATTGATC 60
            ||| |||||||| |||| ||| ||||| ||| |||| |||||||||||||
Db   1308 TAAACATTGAGCTAAATGCGTAAATTCTTCGCCGCGGCGCTTTTCAAAAGAAGCAAACTGATC 1249

Qy   61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGGTGCTTTGACTGGTGCCAA 120
            ||||||||||||| |||||||| ||| ||||| |||||||| |||| ||||||
Db   1248 GAGACTTGCACAAGCAGGCGGCAATPAATACCA----------TATCTCCGCTTTGCAA 1201

Qy   121 TACTGCCCGATCAAGTGTACCAACCTGATATATTATTTAGATGCCAAAGTGCATCACCTTT 180
            ||| || ||| ||| || ||| ||||||||||||||||||
Db   ... 121

Qy   361 GCCGATGTTTGTACCTTTGAATTCATTAAAATATGCCTTGCTATCATAATCATCAATATA 420
Db   975 CCCCACATTTGTTGCTTTAGAGTCATTAATCCAACGAAGTCCATTAGCTTGATGCACTAA 916

Qy   421 TTCACAGCCGATGCGTAGCCCTTTGAAAGTCTTAAGGGGTATCAAGCATGGCCATCCATGGG 480
Db   915 TTGAAAACGATGACTAACCCTTTGAAATGACGAAGTGCCGTACCAAATTGAAATCTAAATT 856

Qy   481 AAGATTGGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
Db   855 AATACCTATAGCTTGTCGTGCCAATGCGTTGCTGCCAAAAGTTCATATATAATTATGCGACC 796

Qy   541 TTTGATTTTTAATTCGATGGATTGGTTAAAAGTTTGTTTTTTTGATGATAATAATGCATACC 600
Db   795 AACCAATGTAGCTTCTTCACAAGGTAAAAAATCGTCAAGTCGTAAGTAAGCTATTGTTGTTG 736

Qy   601 ATCACCATCATCAAATAAATAAAAATCTGC-CGTTTGGTGGCATCTTTAGCATTATGATAATG 659
Db   735 GCCATTTTCAGTTTTTTAGCCAATAATCCGCACTACTTTCCGACAAGAAACGGTATGTTTT 676

Qy   660 TGATAAATGGTGGTGATTGGGGGCAAGCCTTGTCAGTTGTGTGGTCAAGCATTGCTTGGGCTAAAT 719
Db   675 CGCTTGATTTTTCGTTTTTCCCCAAAAGTCAGCC------------- 644

Qy   720 TGGCATCGTCTTGGTAGATGAATGGGGGTTTGGCAATTTTCAAAAATATACGGCAATTTTTGTG 779
Db   643 ---TATCTTCATTGTTCAACACCTACTTTAGCATTATGATAAAATGCGTAAATTTGCTT 587

Qy   780 CCAAATAATCCACACATGCCGATCAAGATGACGATGGAGAGAATTTAAAATTG 839
Db   586 GGCGATAAATCTTCTAAAATCCATATAGGCGATCCATATGATCTTCAGTCACGTTCAAGACAG 527

Qy   840 TTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA 899
Db   526 TCGCTGCCGCGCAGCGTTTTAAGCTATAAGTTGTCTCAAGCTGAAAAACTAGAAAAGCTTCTAGTA 467

Qy   900 CCACCAAATCCATATTTTGGATCGTTAATTCAAGTGCAGGCGTGCCAATATTACCG 959
Db   466 CATAAAAGTTCACAAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAATATATTCCG 410

Qy   960 CGACACCAACAAACATGCCTGCATGTGTTTTGCCATCTCGCCTACTATTAATGTCGTAACACAGTGC 1019
Db   409 CCATACCAACTTTCACACCAGCAGCTTTCGGCCCATTCAATAAACTAAAGTAGTTACGGGTAC 350

Qy   1020 TTTTTGGCCATTGAGCCCTGTGAATGGCGATGAGTTGGTTGGTGTTGCTTGAGTTAGATCTC 1079
Db   349 TTTTTACCATTTGAACCTGTAATCCCCACAATTGGCTTTGTTCGCTCGCGCGGCCA------- 298

Qy   1080 GTGCTTTGAGTGTATCAACAATTGAACAATTGACATCGCTGACTCACAGGAATGCCCTTGAGCTT 1139
Db   297 ------------GAATATTCAATATCGCCGGATTACTTCCACTCCCGCTTTAA 257

Qy   1140 TGGCGGGCGACAAACTACTCGGCCGCGTTCTTGGGTTAATGCCCGGCGGGCTGATAGATTGGATTGA 1199
Db   256 GTGCGGGTTTGAATTTCTGGTGTTTTTACCGCAACCCTGGCGTAATAACAATCATATCGC 197

Qy   1200 CTGAGCTTAAGAGTTCACTGTCTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTT 1259
Db   196 TTTCAAGTAACCATTCCTGATTTAAACTACCACGTATGAAGAGGCATATTTTTGAGGAAGTT 137

Qy   1260 TATCTGCCAAGGTAGGATTGGGGGATTACCATCGGATGCTAACTTTGTGACCGTGATTGA 1319
Db   136 TATCAATACCAGTAGGATTTTTCGGAGTAATCAACTCAACGGAATTAGCCTGTTGGGATA 77

Qy   1320 CCAAAAAATTCACAGCTGACAGTCCCAAACCGACGACGACGGCGATGATAGTTATTTTTT 1377
Db   76 AGAGATAATCCACACAGAAGAACAGACCTGTTTTGCCAAGCCGGATGATAGTAATAATTTTT 19

RESULT 7
US-09-754-468-44/c
; Sequence 44, Application US/09754468
; Publication No. US20050192237A1
; GENERAL INFORMATION:
; APPLICANT: Iversen, Patrick L.
; TITLE OF INVENTION: Antisense Antibacterial Cell Division
; TITLE OF INVENTION: Composition and Method
; FILE REFERENCE: 0450-0033.30
; CURRENT APPLICATION NUMBER: US/09/754,468
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: US 60/174,484
; PRIOR FILING DATE: 2000-01-04
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 11498
; TYPE: DNA
; ORGANISM: Hemophilis influenza
US-09-754-468-44

Query Match   11.6%; Score 160.4; DB 3; Length 11498;
Best Local Similarity   49.6%; Pred. No. 5e-33;
Matches 683; Conservative   0; Mismatches 606; Indels   89; Gaps 7;

(The remainder of this page consists of a pairwise nucleotide sequence alignment presented in rotated orientation, with "Qy" (query) and "Db" (database) rows and flanking base-position numbers spanning positions 1–1377 for the query and 60–9122 for the database.)

```
RESULT 8
US-10-329-670-1/c
Sequence 1, Application US/10329670
Publication No. US20040018503A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
TITLE OF INVENTION: Thereof, and Uses Thereof
FILE REFERENCE: PB186P1
CURRENT APPLICATION NUMBER: US/10/329,670
CURRENT FILING DATE: 2002-12-24
PRIOR APPLICATION NUMBER: US 09/643,990
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: US 08/487,429
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
```

```
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a, t, g or c
```

```
Db   8609  AACCAATGTAGCTTCTTCACAAGGTAAAATCACTCATCTTTACCATTAAATATTGCTT  8550
Qy    601  ATCACCATCATCAAATAAATAAAAATCTGC-CGTTTGGTGGGGCTAAGTAAGCTATTTGTTG  659
Db   8549  GCCATTTTCAGTTTTTAGCCAATAATCCCGCACTATTTTCCGACAAAAGAAACGGGTATGTTTT  8490
Qy    660  TGATAATGGTGTGATTGGGGCAAGCCTTGTCAGTGGTCAAGCAATTGCTTGGCTAAAT  719
Db   8489  CGCTTGATTTTCGTTTCCCCAAAAGTCAGCC------  8458
Qy    720  TGGCATCGTCTTGGTAGATGATGGCGGGTTTGGCAAATTTCAAAAATACGCAATTTTTGTG  779
Db   8457  ---TATCTTCATTGTTCAACACACTTAGCATTATGATAAAATGCGTAAATTTTTGCTT  8401
Qy    780  CCAAATAATTCCACACATGCCATGTGCCGATTCAAGATGATCAGCGCGGAGAGATTTAAAATTG  839
Db   8400  GGCCGATAATCTTCTTAAAATCCATATAGCGGATCCAATTGATCTTCAGTCACGTTCAAGACACAG  8341
Qy    840  TTGCCCCTTGAGCTGCCAAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA  899
Db   8340  TCGCTGCCGCAGCTTTTAAGCTATAAGTTGTCTCAAGCTGAAAAACTAGAAAAGCTTCTAGTA  8281
Qy    900  CCACCAAATCCATATATTTGGATCGTTTAATAATCAAGTGCCAGGCGTGCCAATATTACCCG  959
Db   8280  CATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGGGAATCCCAATATTTCCGC  8224
Qy    960  CGACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAAATGTCGTAACAGTGC  1019
Db   8223  CCAPACCAACTTTCACACAGCAGCTTTCGCCATTTCATAAACTAAAGTAGTTACGGTAC  8164
Qy   1020  TTTTGGCATTGAGCCCTGTGATGGCGATGAGTTGGTGTGTTTGCTTGGAGTTAGATCTC  1079
Db   8163  TTTTACCAATTGAACCCTGTAAATCCCCACAAATTGGCTTTGTCTGCGTCGCGCGGCA-----  8112
Qy   1080  GTGCTTTGAGTGTATCAATGAACAATTGACACATGCCTGCTCACACAGGAATGCCCTTGAGCTT  1139
Db   8111  ----------GAATAATTCAATATGCCGCGGATTTACTTCCCACTCCCCGCTTTAA  8071
Qy   1140  TGGCGGGCGACAATACTCGCGGCGTTCTTCTTGGGTTAANTGCCGGGGCGTGATGAGTGATGAATCGATCGG  1199
Db   8070  GTGCGGGTTTGAAATTTCTGGTGTTTTTACCGGCAAGCCCTGGCTAATAACAATCATATCGC  8011
Qy   1200  CTGAGCTTAAGAGTTCACTGTCTCCAAAGTTCCAAAAATGACAACGCCTGTGGGCAGTT  1259
Db   8010  TTTCAAGTAACCATTCCTGATTTAAACTACCACGTATGAAGAGGGATATTTTGAGGAAGTT  7951
Qy   1260  TATCTGCCAAGGTAGGATTACCATCGATGATGCTAACTTTGTTGACCGTGATTGA  1319
Db   7950  TATCAATACCAGTAGGAGATTTTTTCGAGTATCAATCAACTACACGAATTATTAGCCTGTTGGGAATA  7891
Qy   1320  CCAAAAAATTCACAGCTGACAGTCCCAGAGCTGCCCAAACACCGACGACGCGGCCATATATTTTTT  1377
Db   7890  AGAGATAATTCACACAAGAAAAGACCTGTTTTGCAAGCCCGATGATAGTAATATATTTTTT  7833
```

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)
```

```
Query Match      11.6%;    Score 160.4;   DB 7;   Length 1830121;
Best Local Similarity   49.6%;
Matches 683;   Conservative    0;   Pred. No. 9.3e-32;
                                    Mismatches 606;   Indels 89;   Gaps 7;

Qy         1 TAAAGGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 60
             ||||
Db   1204439 TAAACATTGAGCTAAATGCGTAAAATTCTTCGGCCGCGGCTTTTTCAAAGAGAGCAAAC TGATC 1204380

Qy        61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCCGTGCTTTGACTGCTGCCAA 120
             |  ||||||||||||| |||||| |||| |||| ||
Db   1204379 GAGACTTGCACAAGCAGGCCGACAAATAATACCA                TATCTCGCTTTGCCAA 1204332

Qy       121 TACTGCCCGGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAATGTGCATCACCTTT 180
             | |
Db   1204331 TGTTGGCGGTAAAAATTCTATCGCTTGTTCCATTGTATCGAACCAA     ATAACTTTG 1204278

Qy       181 GAGTAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACTATCCACCCA 240
             ||      ||
Db   1204277 CGATGAAAATTTTGCAAGCAGCGCACCATCTCGACCAAAACAATAACAATAAGTGTGTGG 1204218

Qy       241 CTGCATAACATATGCTGACAGCTCACTCACATAAAATCTTGACTTTACCAAGCCCACCAAAAT 300
             ||
Db   1204217 TTGATTAATTAATTCAGCTAAATTCTGAAAAAATCAGCCCTTTTCCGTCTCCGCCTAGCAA 1204158

Qy       301 CAATGCCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCACTGTGCAATGGTTGA 360
             ||
Db   1204157 CAAATGCAATTTACCCTCAATATA          AAGCCCAGCCAATGCAGCAACTGTACT 1204107

Qy       361 GCCGATGTTTTGTTGCTTTAGGTAAATCATTAAAATATGCCTTGCTATCAAATATCAATATA 420
             ||
Db   1204106 CCCCACATTTGTTTGCTTTAGAGTCATTAATCCACGAAATGCCATAGCTTGATTGCAGACTAA 1204047

Qy       421 TTCACAGCGATGCGGGTAGCCCTTTGAAAGTCTTAAGGGTATCAACGCATGGCATCCATGGG 480
```

RESULT 9
US-10-158-865-1/c
Sequence 1, Application US/10158865
Publication No. US20040203093A1
GENERAL INFORMATION:
APPLICANT: Fleischmann et al.
TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
TITLE OF INVENTION: Thereof, and Uses Thereof
FILE REFERENCE: PB186P2C1D1
CURRENT APPLICATION NUMBER: US/10/158,865

CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US 09/557,884
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: US 08/476,102
PRIOR FILING DATE: 1995-06-07
PRIOR APPLICATION NUMBER: US 08/426,787
PRIOR FILING DATE: 1995-04-21
NUMBER OF SEQ ID NOS: 1
SOFTWARE: Patentin version 3.1

SEQ ID NO 1
LENGTH: 1830121
TYPE: DNA
ORGANISM: Haemophilus influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4747)..(4747)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g

```
Db  1204046 TTGAAAACGATGATCTAACCCTTTGAAATGACGAAGTGCCGTACGAAATTGAAATCTAAATT 1203987
Qy      481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGGCC 540
Db  1203986 AATACCTATAGCTTGTGCCAATGCTGTTGCTGCCAAAATGTTCATTATTATTAATTATGCGGACC 1203927
Qy      541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTTGATGATATAAATGCCATACC 600
Db  1203926 AACCAATGTAGCTTCTTCACAAGGTAAAATCACTTCATCTTTACCATTAAAATATGCTTT 1203867
Qy      601 ATCACCATCATCAAATAAATAAAAATCTGC--CGTTTGGTTGGCTAAGTAAGCTATTTGTTG 659
Db  1203866 GCCATTTTCAGTTTTTAGCCAATAAATCCGCGACTACTTTTCCGACAAAAGAAACGGTA);GTTT 1203807
Qy      660 TGAATAATGGTGTGAATTTGGGGCAACGCTTGTTGGTGGTCAAGCATTGCTTGGCTAAAT 719
Db  1203806 CGCTTGATTTTTCGTTTTCCCAAAAGTCAGCC------------ 1203775
Qy      720 TGGCCATCGTCCTTGGGTAGATGAATGGCGGTTTGGCAAATTTCAAAAATACGGCAATTTTGTG 779
Db  1203774 ---TATCTTCATTGTTCAACACACTACTTTAGCATTAATGATAAAATGCGTAAATTTTGCTT 1203718
Qy      780 CCAAATAATCCACCATGCCGAТCAAGATGACTCAGCGGAGAGATTTAAAATTG 839
Db  1203717 GGGGATAAATCTTCTAAAATCCATATAGCGATCCAATATCTTCAGTCACGTTCAAGACAG 1203658
Qy      840 TTGCCCCTTGAGCTGCCAAATTTGGAGATATGCTCCAAGCTTGAAAACTTGACAGCTCAAGCA 899
Db  1203657 TCGCTGCCCGCCAGCTTTTTAAGCTATTAAGTTGTCTCAAGCTGAAAAACTAGAAAAGCTTCTAGTA 1203598
Qy      900 CCACCAAAATCCATATTTTGGGATCGTTAATAATCCAAGTGCCAGGGGTGCCAAATTTACCGG 959
Db  1203597 CATAAAGTTCACAAATCTTCAATTC---AACAATGACAAAGCGGAATCCCAATATTTCCGC 1203541
Qy      960 CGACACCAACAATCATGCCTGCACGTTTTTGCCATCTCGCTACTAATGTCGTTAACAGTGC 1019
Db  1203540 CCATACCAACTTTCCACCAGCACCTTTCGCCATTCATTTCATAAACTAAAGTAGTTACGGTAC 1203481
Qy     1020 TTTTTGGCATTTGGAGCCCTGTGATGGCCGATGAATTGGTGTTGTTGTTGCTTGAGTTAGATCTC 1079
Db  1203480 TTTTTACCATTTGAACCTGTAATCCCACAATTGCTTTGTCGGCTGGCGGCGCA-------- 1203429
Qy     1080 GTGCTTTGAGTGTATCAATGAACAAATTGAACATCGCTGCTGATCACAGGAATGCCTTGAGCTTT 1139
Db  1203428 ---------------GAATAAATTCAAATATCGCGATTACTTCCCACTCCCGCTTTAA 1203388
Qy     1140 TGGCGGCGCGACAAATACTCCGGCGTTCTTTGGGTTAATGCCGGGGCTTGATGATGATCTGATCG 1199
Db  1203387 GTGCGGGTTTTGAATTTCTGGTGTTGTTTTTACGGCAAGCCTGGTAATAACAATCATTATATGCC 1203328
Qy     1200 CTGAGCTTAAGAGAGTTCACTGTCTCCAAAGCTGCCAAAAATGACGCCTGTGGGGCAGTT 1259
Db  1203327 TTTCAAGTAACCATTCCTGATTTAAAACTACCAGTATGAAGAGGGATATTTTGAGGAAGTT 1203268
Qy     1260 TATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGTGTAACTTTGTGTGACCGTTGA 1319
Db  1203267 TATCAATACCAGTTAGGATTTTTCGGAGTATCAATCAACAGAATATTAGCCTGTTGATTGA 1203208
Qy     1320 CCAAAAAATTCACAGCTGACAGTCCCAAACCGACCACGCCATATTTTTT 1377
Db  1203207 AGAGATAATCACACAGCAAGAAAGACCTGTTTTGCCAAGCCCGATGATAGTAATATTTTT 1203150
```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:

NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)

Query Match   11.6%; Score 160.4; DB 8; Length 1830121;
Best Local Similarity 49.6%; Pred. No. 9.3e-32;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```
Qy        1 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTGCTCATTATATAATTTTTAAATTGATC 60
            ||||    ||| || ||  |||||| | |  |  | || ||    |||| ||||||||||||
Db  1204439 TAAACATTGAGCTAAATGCGTAAAATTCTTCGCCGCGCGCTTTTCAAAAGAAGCAAAACTGATC 1204380

Qy       61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGGTGCTTGACTGGCTGGCAA 120
            ||| || |||||||||| | ||||  |  |  || ||| |||  ||||||||   |||||
Db  1204379 GAGACTTGCACAAGCAGGCGGACAAATAATACCA------------TATCTCCGCTTTGCAA 1204332

Qy      121 TACTGCCCGATCAAGTGTACCAAACCTGATATATTTTAGATGCCAAAATGTGCATCACCTTT 180
            | ||| ||| |  ||| | ||| || || ||||   ||| || ||||  ||||| |||||
Db  1204331 TGTTGGGCGTAAAAAATTCTATCGGTTGTTCCATTGTATCGGAACAA------ATAACTTTG 1204278

Qy      181 GAGTAAATCTTGCTCAATTATCGGGGCCATCTTGACCGGATTAAATACACACTATCCACCCA 240
            | ||||| ||||| |||| |||||| |||||| |||| ||| ||||| |||| ||| || |
Db  1204277 CGATGAAAATTTTGCAAGCACGCACCATCTCGACCAAAAACAATAACAAATAATGTGTGG 1204218

Qy      241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGAACTTACCAAGCCCAACCCAAAAT 300
            ||| || ||| ||||| || | | ||| ||||| |||    |||||  |||| |||||
Db  1204217 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCGGTCTCCGCCCTAGCAA 1204158

Qy      301 CAAATGCCAGTGATGATGATGCCATAAACGTGCCAAGCCCATCAATTGCTGCAAATGGTTGA 360
            |||  |||  ||  ||  ||||| ||||| ||||  |||| || || |||| |||| || ||
```

```
Db  1204157  CAAATGCAATTACCCTCAATATA--------AAGCCCAGCCAATGCAGCAACTGTACT  1204107
Qy      361  GCCCGATGTTGTTGTTGTTACCTTTTGAATCATTAAAAAATATGCCTTGCTATCAAAATATCATCAATATA  420
Db  1204106  CCCCACATTTGTTGTTGCTTTAGAGTCATTAAATCCACAAGGAAATGCCATTAGCTTGATGCACTAA  1204047
Qy      421  TTCACAGCGGATGCGGGTAGCCCCTTTGAAAGTCTTAGGGTATCAAGCACATGGCATCCATGGG  480
Db  1204046  TTGAAAACGATGATCTAACCCTTTGAAAATGACGAAGTGCCGGTACGAAATTGAATCTAAAATT  1203987
Qy      481  AAGATTGGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGACTAGGTTATTGCTTGCC  540
Db  1203986  AATACCTATAGCTTGTGCCAAATGCTGTGCTGCCAAAATGTTTCATATAATTATATGCGACC  1203927
Qy      541  TTTGATTTTTTAATTGAATTGGATTTGGTTAAAAGTTGTTTTTTTTTTGATGAATAAATGCCATACC  600
Db  1203926  AACCAAATGTAGCTTCTTCACAAGGTAAAAAACACTTCATCTTTTTACCATTAAATAATTGCTT  1203867
Qy      601  ATCACCATCATCAAATAAAATAAAATCTGC-CGTTTGGTGGCTAAGTAAGCTATTTGTTG  659
Db  1203866  GCCATTTTCAGTTTTTAGCCAATAATCCGCACTATTTTCCGCAAAAGAAACGGTATGTTTT  1203807
Qy      660  TGATAATGGTTGGTTGATTTGGGGCAACGGCTTGTCTCAGTGTGGTCAAGCATTGCTTGGCTAAAT  719
Db  1203806  CGCTTGATTTTCGTTTTCCCCAAAAGTCAGCC------------------------  1203775
Qy      720  TGGGCATCGTCTTGGTAGAGATGANGGCGGGTTTGGCAATTTCAAAAATACGCGAATTTTTGTG  779
Db  1203774  ---TATCTTCATTGTTCAACACACTACTTTAGCATTATGATAAAATGCGTAAATTTTTGCTT  1203718
Qy      780  CCAAATAATCACCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTG  839
Db  1203717  GGGGATAATCTTCTRAAATCCATATAGGGATCCATATGGATCTTCAGTCACGTTCAAGACAG  1203658
Qy      840  TTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA  899
Db  1203657  TCGCTGCCCGCAGCTTTTTAAGCTATAAGTTGTTCTCAAACCTAGTAGAAAGCTTCTAGTA  1203598
Qy      900  CCACCAAAATCCATATTTTGGATCGTTAATAATTCAAGTGCCAAGGCGTGCCAATATTACCGC  959
Db  1203597  CATAAAAGTTCACAAATCTTCATTC---AACAATGACAAACGGGAAATCCAATAATTTCCGC  1203541
Qy      960  CGACACCAACAACAATCATGCCTGCAGTTTTGCCATCTCGCCTACTAAGTCGGTAACAGTGC  1019
Db  1203540  CCATACCAACTTTCACACACAGCACTTTCGGCCATTCATAAAACTAAAGTAGTTAACGGGTAC  1203481
Qy     1020  TTTTTGGGCATTTGAGGCCTGCATGGCGGATGATTGGGTGTGTTTTGTTCTTCAGTTAGATCTC  1079
Db  1203480  TTTTTACCATTTGAACCTGTAAATCCCCACAATTGGCTTTTGTCGCGTGCGGCA---------  1203429
Qy     1080  GTGCTTTGAGTGTATCAATGAACANACTACTTCACCTTTAA  1139
Db  1203428  --------------GAATAATTCAATATGCCGATTACTTCCACTCCCGCTTTAA  1203388
Qy     1140  TGGCGGCGGACAATACTCGCGGGTTCTTGGGTTAAGCGCTGGGGGCTGAATGATGATCGATCGG  1199
Db  1203387  GTGCGGGTTTGAATTCTCGTGTTTTTACGCGAACCCTGGGCTAATACANCATATCGC  1203328
Qy     1200  CTGAGCTTAAGAGTTCACTGTCCAAAATGACAAATGAACGGCCTGTGTGGGCAGTT  1259
Db  1203327  TTTCAAGTAACCATTCCTGATTTAAAACTACCAGTTATGAGAGGGATTATTTTTGAGGAAGTT  1203268
Qy     1260  TATCTGCCAAGGTAGGATGGGGATTACCATCGATGATGCTAACTTTGTACCGTGATTGA  1319
Db  1203267  TATCAATACCAGTAGGATTTTTCGAGTATCAATCACACAGAATATTAGCTCGTGGGATA  1203208
Qy     1320  CCAAAAAATTCACAGCTGACAGTTCCAGAGCTGCCCAAACCGACGGGACCGATAGTAATAATTTTT  1377
Db  1203207  AGAGATAATCCACACAAGAAAGACCTGTTTTTGCCAAGCCCGATGATGAGTAATATATTTTT  1203150
```

; Sequence 1, Application US/10981687
; Publication No. US2005013122A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Fra
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2CID12
; CURRENT APPLICATION NUMBER: US/10/981,687
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US 10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a, t, c, or g
; FEATURE:
; NAME/KEY: misc_feature

LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117116)..(117116)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)

OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:

```
Query Match
Best Local Similarity   11.6%;   Score 160.4;  DB 9;  Length 1830121;
Matches  683;  Conservative  49.6%;  Pred. No. 9.3e-32;  Indels  89;  Gaps  7;
                          0;  Mismatches 606;

Qy        1 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC   60
              |||     |||||        ||||||||||    ||||||||      |||||||||||
Db  1204439 TAAACATTGAGCTAAATGCGTAAAATTCTTCGCCGCCGCTTTTCAAAAGAAGCAAACTGATC 1204380

Qy       61 CATGCTAGCACAAGCAGAGGTGACAAAAGCACAGCCCTAGGCGGGTGCTTTGACTGGCTGCCAA  120
              ||  |||   |||||||       ||||||||         |||           ||||||
Db  1204379 GAGACTTGCACAAGCAGGCGACAATAATACCA----------TATCTCCGGCTTTGCAA 1204332

Qy      121 TACTGCCCGATCAAGTGTACCAAACCTGATATATTTTAGATCCCAAATGTGCATCACTTT  180
              |||  |||        ||||      |||||    ||||||||  |  |||||||||
Db  1204331 TGTTGGGCGTAAAAAATTCTATCGCTTGTTCCCATTGTATCGAACAA------ATAACTTTG 1204278

Qy      181 GAGTAAAATCTTGCTCAATTATCGGGGGGCATCTTGACCGGATTAAATACACACTCATCCACCCA  240
              |||      |||  |      ||       |||    ||||             ||||||
```

```
Db  1204277  CGATGAAAATTTGCAAGCAGCCACCATCTCGACCAAAAACAATAATGTGTGG                              1204218
Qy      241  CTGCATAACATATGCTGACAGCTCACTAACTTGACCTTTACCAAGCCCACCCAAAAT                         300
Db  1204217  TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTTCCGTCTCCCTAGCAA                       1204158
Qy      301  CAATGCCAGTGATGAATATGCCATAAAACCGTGCCAAGCCCATCAAATTGCTGCAAATGGTTGA                  360
Db  1204157  CAAATGCAATTTACCCTCAATATA----------AAGCCCAGCCCAATGCAGCAACTGTACT                     1204107
Qy      361  GCCGATGGTTTGTACCTTTTGTGAATCATTAAAAAATATGCCATAAAACCGTGCCAAGCCCATCAAATTG            420
Db  1204106  CCCCACATTTGTTGCTTTAGTGAGTCATTAAATCCAATCACACGAATTGGCTTGATAGCACTAA                  1204047
Qy      421  TTCACAGCCGATGGCGGTTAGCCCCTTTGAAAAGTCTTAAGGGTATCAACGATGGCCATCCATCAATAT             480
Db  1204046  TTGAAAACGATGATCTAACCCTTTGAAATGACGAAGTGCGGTTACGAATTGAATCTAAAATT                    1203987
Qy      481  AAGATTGGCAGCTGTGCCCTAAGGCAGAGTGCAGATAAGGCATTGAGTGAGTTAGTGCTTGCC                   540
Db  1203986  AATACCTATAGCTGTGTGCCAATGCTGTTGCCAAAAATGTTCATTATAATTATGCCGACC                      1203927
Qy      541  TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTGATGATAATAATGCCATACC                      600
Db  1203926  AACCAATGTAGCTTCTTCACAAGGTAAAAATCACTTCATCTTTTACCCATTAAAATATTGCTT                   1203867
Qy      601  ATCACCATCATCAAAATAAAAAATCTGC-CGTTTGGTTGGCTAAGTAAGCTATTTGTTG                       659
Db  1203866  GCCAATTTTCAGTTTTTAGCCAATAATCCGCACTATTTTCCGCAAAAGAAACGGTAATGTTT                    1203807
Qy      660  TGATAAATGGTGTGATTGGGGCAACGCCTTGTAGTGGTGCAAGCATTGCTTGGGCTAAAT                      719
Db  1203806  CGCTTTGATTTTCGTTTTTCCCCAAAAGTCAGCC-----------                                     1203775
Qy      720  TGGCATCGTCTTGGTAGATGATGGGGCGGTTTGGCAATTTCAAAAAATACGCAAT-TTTGTG                    779
Db  1203774  --TATCTTCATTGTTCAACACACCTACTTTAGCATTATGATAAATGCGTAAATTTTGCTT                      1203718
Qy      780  CCAAATAAATCCACCACCATGCGCCGGATCAAGATGACAGCGGAGAGATTTAAAATTG                        839
Db  1203717  GGCGGATAAATCTTCTAAAATCCAATATAGGGATCCATTGCTTCAGTCACCTTCAAGACAG                     1203658
Qy      840  TTGCCCCTTTTGAGCTCGCCAAATTTGAGATATATGCTCAAGCTCAAGCTCAAGCA                          899
Db  1203657  TCCGTCGCCGCAGCGTTTTAAGCTATAAGTTGTCTCCAAGCTGAAAACTAGAAAGCTCTAGTTA                  1203598
Qy      900  CCACCAAATCCATATTTTGGATTCGTTAATAATTCAAGTGCAGGGCCTGCCATCTTTTACCGC                   959
Db  1203597  CATAAAGGATAAATCTTCACAAATCTTCATTC----AACAATGACAAAAGCGGGAATCCCAAATAATTCCGC          1203541
Qy      960  CGACACCAACAAATCATGCCTCTGCAGTGTTTTGCCATCTCGGCTACTAATGTGTCGTAACAGTGC               1019
Db  1203540  CCATACCAAACTTTCACCAGCAGTTTCGTTGTTCGCGTTGAGTCGCGGCA----------                      1203481
Qy     1020  TTTTGGGCCATTGAGTGTATCAATGAACAAATTGAACCTGGTGGCGATGATTGGGTGTTGTTGCTTGAGTTAGATCTC   1079
Db  1203480  TTTTACCCATTTGAACCTGTAATCCCCACAATTGGCTTGTGGCCAAAATTACTTCATCTTTTTACCATTA            1203429
Qy     1080  GTGCTTTGAGTGTATCAATGAACAAATTGGATGAGTTGTGGCCAGGAATGTGCCCTTGAGCTT                   1139
Db  1203428  ------------GAATAATTCAATATATGCCGATTACTTCCACTCTGATTCGG                             1203388
Qy     1140  TGGCGGCGGACAAATACTCGCGGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTTGATCGG                   1199
Db  1203387  GTGCGGTTTGAATTTCTGTTGTTTTTACCGGCATTTATTACAATCATATCGC                             1203328
Qy     1200  CTGAGCTTAAGAGTTCACTGTCTGCCAAAATGACAATGACGCTGCTGTGGGCAGTT                          1259
Db  1203327  TTTCAAGTAACCAATTCCTGATTTAAACTACCAGTATGAAGAGGGATATTTTGAGGAAGTT                     1203268
Qy     1260  TATCTGCCAAAGGTAGGAGTAGGATTTTTGGGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGA         1319
Db  1203267  TATCAATACCAGTAGGGATTTTTGGAGTATCAAATCCATATAGCGGATCATATGATGTTCAGTCACGTT            1203208
```

```
Qy     1320  CCAAAAAAATTCACAGCTGACAGCGTTCCAGAGCGTGCCCAAAACGGACGACGGCGACGCATATTTTTT             1377
Db  1203207  AGAGTAAATCACACAAGAAAGACCGTGTTTGCCAAGCCCGATAGTAGTAAATATTTTT                        1203150

Qy      172  ATCACCCTTTGAGTGTAAAATCTTCGCTCAATTATGCTGGGGGCATCTTGACCGATTAAATAACACACT             231
Db     5280  ATAGCTTTTGTTGGATTGAGAAAAATTTTGCAAGCTGCACACCAATCTCGACCAAAAACAATAACAAAT             5339
Qy      232  ATCCACCCACTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCC                      291
Db     5340  AATGTGTGGGTTGATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTTCCGTCTCC                         5399
Qy      292  ACCCAAAATCAATGCCAGTGATGAATAATTAATTCCATAAAACCGTGCCAAGCCCATCAAATTGCTGC              351
Db     5400  GCCTAGCAACAAAATGCAAT--------TTACCCTCAATATAAAAGGCCCAGCTAATGCAGC                     5450
Qy      352  AATGGTTGAGCCGGATGTTTGTACCTTTTGTGAATCATTAAATAAAATAATGCTTGCTATCAAATATC              411
Db     5451  AACTGTACTCCCCACACATTGTTGCTTAGTGAGTCATTAAATCCAACAGCTTGATAGCACTAA                   5510
Qy      412  ATCAATATATTCACAGCGCGATGCGGGTAGCCCCTTTGAAAAGTCTTAAGGGTATCAAGCATGGC                  471
Db     5511  ATGCACTAATTGAAAACGATGATCTAACCCTTTGAAATGACGAAGTGCGGTTACGAATTGA                     5570
Qy      472  ATCCATGGGAGAGATTGGCAGCTGTGCCTAAGGCAGAGATGCAAGTGCAGCATTGAGTAGGTT                   531
Db     5571  ATCTAAAATTAATACCGACAGTTACTTCTTCACAAGGCAAAAATACTTACTTCATCTTTTTACCATTAA            5630
Qy      532  ATGCTTGCCTTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTGATGATAATGA                      591
Db     5631  AATGGCGACCGGACTAATGTTACTTCTTCACAAGGCAAAATTACTTACTTCATCTTTTTACCATTAA              5690
Qy      592  TGCCATACCATCACCATCATCAAAATAAAATAAATCTGCCGTTTGGTTGGCTAAGTAAGCT                      651
Db     5691  GTATTGCTTGCCAATTTCAGTTTTTCAGTTTTTAGCCAATAATCCGCACTATTTTCCGCAAAAGAAAC             5750
Qy      652  ATTTGTTGTGATAAATGGTGTGATTGGGGCAACGCCTTGTAGTGGTCAAGCATTGCTT                        711
Db     5751  GGTTTGTCTCGCTTGATTTCGTTTTTTCACCTTCACCAAAAGTCAGCC-----------                       5790
Qy      712  GGCTAAATTGGCCATCGTCTTCTTGGTAGATGATGGGGGCGGTTTGGCAATTTCAAAAAATACGCAA               771
Db     5791  ------------TTATCCTCATTATTCAAAACTCCAGCATTATGATAAATCGGTAA                           5840
Qy      772  TTTTGTGTGTCCAAATAAATCCATTCTTCTAAAATCCATGGGATCGATCATAGATATCTTCAGTCACGTT            831
Db     5841  TTTTGCTTGGCCGATAAATCTTCTAAAATCCAATATAGCGGATCATATGATGTTCAGTCACGTT                  5900
```

```
Qy   832 TAAAATTGTTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAG 891
         |||||||||||||| |||||              |||||||| | |||||||||||| |||
Db  5901 CAAAACAGTCCCAGCCGCGAGCTTTTAAGCTATAAGTTGTCTCAAGCTGAAAACTAGAAAG 5960

Qy   892 CTCAAGCACCACCAAATCCATATTTTGGATACGTTAATAATCAAGTGCAGGCGTGCCAAT 951
         |||| |  |||||     |  ||| | |    |  ||||         ||| | | ||||
Db  5961 CTCTAGTACATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAAT 6017

Qy   952 ATTACCGCGACACCAACAATCATGCCTGCAGTGTTTTGCATCTCCCTACTAATGTCGT 1011
         |||| ||| |||||||||  |   |     |           |||  | |  |||||||
Db  6018 ATTTCCGCCCCATACCAACACTTTCACGCGCCAGCAGCTTTTCCATTTAAAACTAAAGTAGT 6077

Qy  1012 AACAGTGACTGGCATTTGACCTGTGGATCGGATGATGAAATGATTGGGTGTGTTGTTGTTTGAGT 1071
         ||||||| | | |||||||     | | || | || || | ||| |  | | ||| | || |
Db  6078 AACAGTACTTTTTTCCATTTGAACCTGTAAATCCCCACAATTGGCTTCATTTCGGTCGCGG--- 6134

Qy  1072 TAGATCTCGTGCTTTGAGTGTATCAATGAACATGCTGATCCTGATCCAGGAAGTGCC 1131
Db  6135 ------------GCAGAATAATTCAATATCGCGATTACTTCCACTCC 6170

Qy  1132 TTGAGCTTTGGCGGGGACAATACTCGGCGTTCTTGGGTTAATGCTCCGGGCGTGATGAGT 1191
Db  6171 CGCTTTAAGTGCGGTTTGAATTTCTGGGTGTTTTACCGGCAACCAAATGCAAT 6230

Qy  1192 CTGATCGGCTGAGCTTAAGAGTTCACTGTCTCCAAGCTGCCAAAAATGACAATGAACGCCTGT 1251
Db  6231 CATATCGCTTTCAAGTAACCATTCCTGATTTAAACTACCAGTATGTAGAGAGGGAATATTTTG 6290

Qy  1252 GGGCAGTTTATCTGCCAAGGTAGGAATTGGGAATTACCATCGATGATGCTACTTTGTGACC 1311
Db  6291 AGGAAGTTTATCAATACCAGTAGGCTTTTTTTCGGGTATCAATCACACGGAATATTAGCCTG 6350

Qy  1312 GTGATTGACCAAAAAATTCACAGCTGACAGTCCCCAAAACCGAGACGACGGCATA 1371
Db  6351 TTGGGATTAAGAGATAATCCACACAACAAGAAGACCTGTTTTGTTGCCAAGCCCAATGATCGTAAT 6410

Qy  1372 TTTTTT 1377
         ||||||
Db  6411 ATTTTT 6416
```

```
Query Match    11.2%;  Score 154.4;  DB 9;  Length 908766;
Best Local Similarity   49.0%;  Pred. No. 2.9e-30;
Matches  591;  Conservative  0;  Mismatches 546;  Indels 69;  Gaps 4;

Qy   172 ATCACCTTTGAGTAAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACT 231
Db 107917 ATAGCTTTGTTGATGAAAATTTTGCAAGCTGCACACCAATCTCGACCAAAAACAATAACAAAT 107858

Qy   232 ATCCACCCACTGCATAAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTTACCAAGCCC 291
Db 107857 AAATGTGGTTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCCGTCTCC 107798

Qy   292 ACCCAAAATCAATGCCAGTGATGAATATGCCATAAACCGTGCCAAGCCCATCAAATTGCTGC 351
Db 107797 GCCTAGCAAACAAAATGCAAT----------TTACCCTCAATATATAAAGCAGC 107747

Qy   352 AATGGTTGAGCCGATGTTGTTGTACCTTTTGAATCATTAAAAATATGCCTATCAATATC 411
Db 107746 AACTGTTACTCCCCACATTTGTTTGCTTTAGAGTCATTAATATCCAACGAATACCATTAGCTTG 107687

Qy   412 ATCAATATATTCACACGATCGGGTAGCCCTTTGAAAGTCTTAAGGTATCAAGCATGGC 471
Db 107686 ATGCACTAATTGAAAACGATGAATCTAACCCCTTTGAAAGAGAAGGTGACGAAGTACGAAATTGA 107627

Qy   472 ATCCAAGGGAAGAGATTGGCAGCTGTGCCTAAGGCAAGTGCGAGATAAGGCAATTAGCTTG 531
Db 107626 ACTAAATTAATACCGACAGCTTGTGTGCCAAATGCTGTTGCTGCCAAAATGTTCATATAATT 107567

Qy   532 ATGCCTTGCCTTTGATTTTTAAATTGATGAATGGGTAAAAGTTGTTTTTTTGATGATATAA 591
Db 107566 ATGGCGACCGGACTAAGTTACTTCTTCACAAGAGGCAAAAATTACTTCATCTTTTTACCATTAA 107507

Qy   592 TGCCATACCATCACCATCACTCAAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCT 651
Db 107506 GTATTGCTTGCCCATTTTCAGTTTTTTAGCCAATAATCCGGCAAAAAGAAAC 107447

Qy   652 ATTTGTTGTGATAAATGGTGTGATTGGGGCAACGCTTGTTGCAGTGTGGTCAAGCCATTGCTT 711
Db 107446 GGTTTGTCTCCGTTGATTTTCACCTTCACCAAAAGTCAGC---------- 107407

Qy   712 GGCTAAATTGGCATCGTTCTTGGTAGATGAGATGGCGGTTTTGGGCAAATTTTCAAAAATAGGCCA 771
Db 107406 ----------TTATCCTCATTATTCAAAAACTCCGACCTCAGCCATTATGATGAAATGGCTAA 107357

Qy   772 TTTTTTGTTGCCAAAATTAACCACCATGCCCATCAAGATGATCAGCGGGAGAGATT 831
Db 107356 TTTTTGGTTTGGGGATAATCCTTAAACTCCATATAGGCGGATCCAATATGGGATCTTCAGTCACGT 107297

Qy   832 TAAAAATTGTTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAG 891
Db 107296 CAAAACAGTCCGCAGCCGCGAGCTTTTAAGCTATAAGTTGTCTCAAGCTGAAAACTAGAAAG 107237

Qy   892 CTCAAGCACCACCAAATCCATATTTTGGATACGTTAATAATCAAGTGCAGGCGTGCCAAT 951
Db 107236 CTCTAGTACATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAAT 107180

Qy   952 ATTACCGCGACACCAACAATCATGCCTGCAGTGTTTTGCATCTCCCTACTAATGTCGT 1011
Db 107179 ATTTCCGCCCCATACCAACACTTTCACGCGCCAGCAGCTTTTCCATTTAAAACTAAAGTAGT 107120

Qy  1012 AACAGTGACTGGCATTTGACCTGTGGATCGGATGATGAAATGATTGGGTGTGTTGTTGTTTGAGT 1071
Db 107119 AACAGTACTTTTTTCCATTTGAACCTGTAAATCCCCACAATTGGCTTCATTTCGGTCGCGG--- 107063
```

RESULT 13
US-10-282-122A-25165/c
Sequence 25165, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yanamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: Patentin version 3.1
SEQ ID NO 25165
LENGTH: 1281
TYPE: DNA
ORGANISM: Legionella pneumophila

US-10-282-122A-25165

Query Match   11.1%; Score 152.8; DB 7; Length 1281;
Best Local Similarity   51.6%;
Matches 464; Conservative 0; Mismatches 412; Indels 24; Gaps 4;

```
Qy   170  GCATCACCTTTGAGTAGAAATCTTGCTCAAATTATCGGGGGCATCTTGACCGATTAAAATACACA  229
Db  1163  GGAACTACTTTGGCTTAACAGCCGGATTCAAATCTTCACCAATCAAAAACAATG          1104

Qy   230  CTATCCACCCACTGCATAACATATGCTAGACAGCTCACTAAAATCTTGACCTTTACCAAGC     289
Db  1103  GAAACGAACAAAATTCTCAAAACTGGCGTGGGCTAATTCTTGGAAAATCCGCCCCTTTTCCCTGT 1044

Qy   290  CCACCCAAAATCAATCGCCAGTGATGATAATCCATAAACCGTACCAAGCCCATCAATTGCT     349
Db  1043  CCACCCGCAATAAGAACAATTTCCTTGAAATCATTAAAATATGCCTTGCTATCAATA       987

Qy   350  GCAATGGTTGAGCGGATGTTTGTTGTACCTTTTGAAATCATTAATCATTAAAAAACC---TCCTATTCCATTAATAGCA  409
Db   986  GAGAGTGTAGCACCAAATATTTGTTCCTTCGAAATCATTAATCCAACCTACTCCATCTACT    927

Qy   410  TCATCAATATATTCACAGCGATGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATG  469
Db   926  TCTCCTTACCCACTGGGCATCCGATTCGATTGCCAGTCCAGGGGAAGGTTTTTAAAACATTAAGAAATG 867

Qy   470  GCATCCATGGGAAGAGATTGGCGAGCTCGTGTCCTAAGGCAGGTGGCTAAGAGTGCAGATAAGGGCATTGAGTAGG 529
Db   866  TGTTGCATAGAAATTCCCGCTGCTTCGGCTAAAGCACAAGCAGCAAGCGCGAATTCATCCAA    807

Qy   530  TTATGCTTGCCTTTGATTTTAATTGATGATTGGTTAAAGTTTGTTTTTTGATGATAT        589
Db   806  TTATGCACCCCTTTTAATTAAAAATCGGATTCCACTGGGCAAAAGAGGCTTCCATTCCTTTCGCC 747

Qy   590  AATGCCATACCATCACCATCATCAAATAAAATAAAAAAATCTCCCGGTTGTAAAGTTGGCTAAGTAAG 649
Db   746  AAATAAGTAGTATTCTCTTGTTCAATCAGCCCCCAAATTTCCCATT------GAAG        698

Qy   650  CTATTTGTTGTTGATTAATGGTGTGATTTGGGGCAACGCCTTGTTCAGTGTGGTCAAGCAATGC 709
Db   697  GCGCGTCTTTGCCAAAAGAGGATGCATTTTTATATCTGGATGGGCGGTTTGGCAATTTTCAAAAAATACGC 642

Qy   710  TTGGGCTAAATTGGCATCGTCTTTGGGTAGGATGAGGGGCGCAATTTTCAAAAATAATCCTT    769
Db   641  ---ACAGTGTAGAGCGTCTTCACGATTAAAATAAAACAGCCTTGGCACCACGGATAAAATCCTT  585

Qy   770  AATTTTTGTTGCCAAATAATCCACCATGCCATCGTTCCGATCAAGATGATCAGCGGAGGA     829
Db   584  TGTTTTTGCTTGGCGTAAGGCTTCCATAGTGTGATCCTGGTCGAAATGGGGTTACA        525

Qy   830  TTTAAAAATTGTTTGCCCCTTGAGCTGCCAAAATTGAGATAATGCTCAAGCTGAAAAAACTTGAC 889
Db   524  TTAAGAATCGTCGAACCACTGGGCACAAGAATAAGTTAAATCAATGAAAACTGGAT        465

Qy   890  AGCTCAAGCACCACCAAAATC---CATATTTGGGATCGTTAATAATTCAAGTGCAGGCGTG     946
Db   464  AACTCCAAATACCAAAATCGTAGTGAGTTCATCATCAACATATCTAAGACAGGAGTG       405

Qy   947  CCAAATATTACCGCCGACACAACAATCATCGGTCCGTGCATGTTTTGCCATCTCGCCTACTAAT 1006
Db   404  CCGATATTTCCGGCTACAGCTACGGGAAAACCCGGCGCCGTTAGCCCATTCTGCCAAA      345

Qy  1007  GTCGTTAACAGTGCTTTTGCCCCTTGATGGCCGATGAGGATGATTGGTGTGTTTTGTTGCT   1066
Db   344  GTCGTTACAGTGGGATTTACCATTAGTTCCGTTATGGCAATGACAGGGGCGGCTAATCTCT    285
```

```
Qy  1072  TAGATCTCGTCGTTTGAGTGTATCAATGAACAACATCGCTTGATCACAGGAATGCC        1131
Db 107062  ----------------GCAGAATAATTCAATATCGCCGGATTACTTCCACTCC          107027

Qy  1132  TTGAGCTTTGCGGCGGACAATACTCGGGCGTTCTTGGGTTAATGCCGGGGGCTGATGATGAT   1191
Db 107026  CGCTTTAAGTGCGCGGTTTGAATTTCTCGGTGTTTTTACCGGAAGCCCCTGGGCTGAATAACAAT 106967

Qy  1192  CTGATCGGCGCTGAGCTTAGAGTTCACTGTCCCAAGCTGCCAAAATGACAATGAACGCCCTGT   1251
Db 106966  CATATCCGCTTTCAGCTAACCATTCCTGATTTAAACTACCAGTATGAAGAGGGAGTATTTTG   106907

Qy  1252  GGGCAGTTTATCTGCCAAGGTAGGAGGATTGGGATTACCATCGGATGATGCGTAACTTGTTGACC 1311
Db 106906  AGGAAGTTTATCAATACACCAGTAGGCCTTTTTTCGGGTATCAATCACGACGAATATTAGCCTG 106847

Qy  1312  GTGATTGACCAAAAAAAATTCACAGCTGACAGTCCCAAAACCGACGACGGCCATA          1371
Db 106846  TTGGGATAAGAGATAATCCACACAAGAAAAAGACCTGTTTTGCCAAGCCCAATGATCGTAAT   106787

Qy  1372  TTTTTTT  1377
Db 106786  ATTTTT   106781
```

RESULT 14
US-10-958-216-402/c
Sequence 402, Application US/10958216
Publication No. US20050181388A1
GENERAL INFORMATION:
APPLICANT: EDWARDS, ALED

```
; APPLICANT:    DHARAMSI, AKIL
; APPLICANT:    VEDADI, MASOUD
; APPLICANT:    ALAM, MUHAMMAD ZAHOOR
; APPLICANT:    ARROWSMITH, CHERYL
; APPLICANT:    AWREY, DONALD E.
; APPLICANT:    BEATTIE, BRYAN
; APPLICANT:    BUZADZIJA, KRISTINA
; APPLICANT:    CANADIEN, VERONICA
; APPLICANT:    DOMAGALA, MEGAN
; APPLICANT:    HOUSTON, SIMON
; APPLICANT:    KANAGARAJAH, DHUSHY
; APPLICANT:    LI, QIN
; APPLICANT:    MANSOURY, KAMRAN
; APPLICANT:    MCDONALD, MERRY-LYNN
; APPLICANT:    NETHERY, KATHLEEN
; APPLICANT:    NG, IVY
; APPLICANT:    OUYANG, HUI
; APPLICANT:    PINDER, BENJAMIN
; APPLICANT:    RICHARDS, DAWN
; APPLICANT:    TAI, MATTHEW
; APPLICANT:    THALAKADA, ROSANNE
; APPLICANT:    VALLEE, FRANCOIS
; APPLICANT:    VIRAG, CRISTINA
; TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM BACTERIA
; FILE REFERENCE: IPT-205.01
; CURRENT APPLICATION NUMBER: US/10/958,216
; CURRENT FILING DATE: 2004-10-04
; PRIOR APPLICATION NUMBER: PCT/CA03/00462
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/369,511
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 60/385,089
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/385,751
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/386,553
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/386,577
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/386,367
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/386,566
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/386,390
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/386,601
; PRIOR FILING DATE: 2002-06-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1132
; SOFTWARE: Patentin Ver. 3.3
;
; SEQ ID NO 402
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-10-958-216-402

Query Match    11.0%;  Score 152.4;  DB 9;  Length 1314;
Best Local Similarity   49.2%;  Pred. No. 2.5e-31;
Matches  678;  Conservative   0;  Mismatches  611;  Indels  89;  Gaps   7;

Qy      1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATATAATTTTAAATTGATC 60
            |||   || ||         ||||   ||| ||  || || ||| ||  |   || |
Db   1308 TAAACATTGAGCTAAAATGCGTAAAATTCTTGCCGCCCGGCTTTTCAAAAAGAACAAACTGATGATC 1249

Qy     61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 120
            || ||| |||||   ||| ||| |  |||   ||||    || ||| |||||  ||||
Db   1248 GAGACTTGCACAAGCAGGGACAATAAATACCA-------TATCTCCGCTTGCAA 1201

Qy    121 TACTGCCCGATCAAGGTACCAACCTGATATATTTTTAGATGCCAAATGTGCATCACCTTT 180
             || ||| |||||  ||||   |||||    | |||||  |||   |||  ||| ||||||
Db   1200 CTGAGCTTAAGAGTTCACTGGTCCAAGCTGCCAAAATGACAAATGAACGCCTGTGGGCAGTT 1259

Qy    181 GAGTAAATCTTGCTCAAATTATCGGGGGCATCTTGACCGATTAAATACACTACTATCCACCCA 240
            |||                                            ||  ||| ||||
Db    196 TTTCAAGTAACCATTCCTGATTTAAACTACCAGTATGAAGAGGAGAGTATCTTGAGGAAGTT 137

Db   1146 CGATGAAAATTTTGCAAGCAGGCGCACCCATTCTGGACCAAAAACAATAACAAATAAGTGTGTGG 1087

Qy    241 CTGCATAACATATGTGACAGCAGCTCACTAACTCTAAAAATCTTGACCTTTACCAAGCCCACCAAAAT 300
                | || | |  ||| |   |  |    |||                  | ||||||  |||||
Db   1086 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCCGTCTCCCGGCTAGCAA 1027

Qy    301 CAAATGCCAGTGATGAATATGCCATAAAACGGTGCCAAGCCCATCAAATTGCTGCAAGGGTTGA 360
            |||  |||| | ||||   |  | ||||||||||||  |||  ||||   || |||   |||
Db   1026 CAAATGCAATTTACCCTCAATATA--------AAGCCCAGCACCCAAATGCAGCAACTGGTACT 976

Qy    361 GCCGATGCGTTTGTACCTTTTTGAATCATTAAAAATATGCCTTGCTATCATCATCAATATA 420
            || |||| || |||  |||| | ||| || ||          |||||| ||| || ||||
Db    975 CCCCACATTGTTGTTCTTTAGAGTGTCATTAATCCAACGGAATGCCATTACCTTGATGCACTAA 916

Qy    421 TTCACAGCCGATGCGGGTTAGCCCCTTTGAAAGTCTTAACGGTATCAAGCATGGCATCCATGGG 480
            ||||  | |  || ||| ||   ||||||  |||||  | |||  ||  ||| | || |||
Db    915 TTGAAAACGATCATCTAACCCTTTGAAATGAGCAGTGCGGTACGAATTCTAAATT 856

Qy    481 AAGATTGGCCAGCTGCCTAAGGCAAGTGCAGATAAGGCAATTGAGTAGGGTTATGCCTTGCC 540
            |||   || ||| || | |||  | || ||||  |||| ||| ||| || |||||| |||
Db    855 AATACCTATAGCTTGTGGCCAATGCTGTTGCTGCGCCAAAAGTCAGCC 796

Qy    541 TTTGATTTTAATTGATGGATTGGTAAAAGTTGTTGTTTTTTGCATGATAATAATGCCATACC 600
            ||||||                 |||||  |||||||||| ||| |||||||||  ||
Db    795 AACCAATGTAGCTTCTTCACAAAGGTAAAAATCACTTTTACCAATTAAAATATGCTT 736

Qy    601 ATCACCATCATCACAAATAAAATAAAAATCTGC-CGTTTGGTGGCTAAGTAAGCTATTTGCTTG 659
            ||||  | || |   |||||||   | || |||  |||||| |||||| ||| ||||||||||
Db    735 GCCATTTTCAGTTTTTTTAGCCAATAATCGCCGCCTACTATTTTCCGAAAAGAAACCGGTATCGTTT 676

Qy    660 CGCTTGATTTTCGTTTTCCCCAAAAGTCAGCC-- 644
            |||||| |||  |||| ||  |||| ||||||
Db    675 TGATAATGGGTGTGCATTGGGGGCAACGCTTGTCAGTGTGGTCAAGCTTGGCTAAAT 719

Qy    720 TGGCATCGTCTTGGTAGATGATGGCGGTTTGCAATTTTTCAAAAAATACGGCAATTTTGTG 779
                 |||| || ||| | | || ||||   ||       |||||||||       |||||
Db    735 GCCATTTTCAGTTTTTTTAGCCAATAATCGCCGCCTACTATTTTCCGAAAAGAAACCGGTATCGTTT 676

Qy    643 ---TATCTTCATTGTTCAAACACACTACTTTCATCTTTAGCATTATGATAAATTTTGCTT 587
                |||||||| |||    |||| ||||  | |||  |||| ||| ||  | ||||||
Db    644

Qy    780 CCAAAATAAATCCACCATGCCATCGTGCCGATCCAAGATGATCAGCGGCGGAGGATTTAAAATTG 839
            |||||| |  ||||||  ||| | ||  |||| || ||||||||| | || ||| | || |||
Db    735 GCCATTTTCAGTTTTTTTAGCCAATAATCGCCGCCTACTATTTTCCGAAAAGAAACCGGTATCGTTT 676

Qy    586 GGCGATAATCTTTCTAAAATCCATTAAGCGATCCATAGGATTTCAGTGTACGGTTCAAGACAG 527
            |||  |||| ||| || |||||  |   ||| ||| ||| ||| |||| | ||| |||  |||

Qy    840 TTGCCCCTTGAGCGTGCCAAAATTTGAGATATGCTCAAGCGAAAAACTTGACAGCTCAAGCA 899
            ||| || |||| | || |||||||||| || || || ||| || || || || || ||||
Db    899

Qy    526 TCGCTGCGCCGCAGCTTTTAAGCTATAAGCTGTCTCAAGCTGAAAAGCTCTTTTAAGTA 467

Qy    900 CCACCAAAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGCGGTGCCCATATTATTACGGC 959
            ||| ||||| |||| ||  |||||| || || ||||||||| | ||| |||||    || ||||
Db    467

Qy    466 CATAAAGTTCACAATCCTCATTC---AGCAATGACAAAAGCGGGAATCCCAATTATTTCCGC 410

Qy    960 CGACACACAACAATCAATGCCTGCAATCATCGCCTACTTACTGCCTTTGCCCTACCTAAATGTCGTAACAGTGC 1019
            |||| |||  |||| |   ||| || |   |||| |   || || | ||||||||||| | |||||| | |
Db    409 CCATACCAACTCTAACTCTACGAGTTCGCGCTGTTTCATTAAACTAAAGTAAGTAGTTACGGTAC 350

Qy   1020 TTTTTGGCATTTGAGCCTGTGATGGCGGAATGGATTGGGTGTGTTGTTGCTTGAGATCTC 1079
            |||||                    ||||       || ||  | |||| |||| ||
Db    349 TTTTTACCATTAGATCCTGTAATCGTTAATAATCCCCACAATTGGCTTTTGTCGCCTGCCGGCA----- 298

Qy   1080 GTGCTTTGAGTGTATCAATGAACAATCGCTGATCACAGAGAATGCCTTTGAGCTT 1139
            |||||| ||       |||  ||   | || || ||  |||  |||| | |||||
Db    297 ------------------GAATAAATTCAATATCGGCGGGATTTACTTCCACTCCGCTTTAA 257

Qy   1140 TGGGCGGGCGACAAATACTCGCGGCGTTCTTGGGGTTAATGCGGGCGATGATGGAGTACTTCGGCGG 1199
            || ||     ||||| |  |  ||| || |  ||| ||||  ||| ||||| || |||| | |||
Db    256 GTGCGGTTTGAATTTGTGTTTTTTTACCGCAAGCCCTAATAACAATCATTACGC 197

Qy   1200 CTGAGCTTAAGAGTTCACTGGTCCAAGCTGCCAAAATGACAAATGAACGCCTGTGGGCAGTT 1259
            |  |||||| |||||||| || || ||||||||||| | || ||| |||||| | |||||||
Db   1200 CTGAGCTTAAGAGTTCACTGGTCCAAGCTGCCAAAATGACAAATGAACGCCTGTGGGCAGTT 1259

Qy    196 TTTCAAGTAACCATTCCTGATTTAAACTACCAGTATGAAGAGGAGAGTATCTTGAGGAAGTT 137
            |   |||                               |||  ||||||   || ||||||

Qy   1260 TATCTCCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGA 1319
            ||||                               ||||| || |||  || |||||||
```

```
RESULT 15
US-10-282-122A-30713/c
; Sequence 30713, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Liangsu
;  APPLICANT: Zamudio, Carlos
;  APPLICANT: Malone, Cheryl
;  APPLICANT: Haselbeck, Robert
;  APPLICANT: Ohlsen, Karl
;  APPLICANT: Zyskind, Judith
;  APPLICANT: Wall, Daniel
;  APPLICANT: Trawick, John
;  APPLICANT: Carr, Grant
;  APPLICANT: Yamamoto, Robert
;  APPLICANT: Forsyth, R.
;  APPLICANT: Xu, H.
;  TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
;  FILE REFERENCE: ELITRA.034A
;  CURRENT APPLICATION NUMBER: US/10/282,122A
;  CURRENT FILING DATE: 2003-02-20
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/230,335
;  PRIOR FILING DATE: 2000-09-06
;  PRIOR APPLICATION NUMBER: 60/230,347
;  PRIOR FILING DATE: 2000-09-09
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/267,636
;  PRIOR FILING DATE: 2001-02-09
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 30713
;  LENGTH: 1305
;  TYPE: DNA
;  ORGANISM: Pasteurella multocida
US-10-282-122A-30713

Query Match            10.3%;  Score 142.2;  DB 7;  Length 1305;
Best Local Similarity  48.8%;  Pred. No. 1.7e-28;
Matches  577;  Conservative  0;  Mismatches 528;  Indels 78;  Gaps  4;
```

```
Qy     136 TATCAAATACCAGCAGGAGATTTTTTCGAGTATCAATCACACGAATAATTAGCCTGGTGGGATA 77
               ||| ||| |||| | |||||| | |||| ||  |   ||  |   |||   ||| |
Db    1320 CCAAAAAATTCACAGCTGACGTCCAGAGCTGCCAAAACCGACGACGGCCAATATTTTTTT 1377

Db      76 AGAGATAAATCCACAAGAAAGAACCTGTGTTTTGCCAAGCCGGATGATAGTAATATTTTTT 19
```

```
Qy     195 CAATTATCCGGGGCACTCTTGACCGGATTAAATACACACTACCCACCACTCCATAACATATG 254
Db    1126 CAAGTTGGTTCGCCATTCTTGACCAAAAACAATAGCAAAATATATTCGGTTGATTAATTAAAC 1067

Qy     255 CTGACAGCTCACTAAAAAATCTTGACCTTTACCAAGCCCCAACCCCAAAAATCAATGCCAGTGATG 314
Db    1066 TGGCTAATTCCGAAAAGTCCGCACCTTTACCATCACCACGAGTAACAAAATGCAAATGTCC 1007

Qy     315 ATATGCCATAAAACCGTGCCAAGCCCATCAAATTGCTGCAATGGTTGTTGAGCCGGATGTTTGTAC 374
Db    1006 CGTTTAACT----------GTAAGCCAGTTAATGCGGGCAACAGTACTGCCCACATTGGTGG 956

Qy     375 CTTTTGAATCATTAAAAATATGCCTTGCTATCATAATATCATAATATATTCAGCGATGCG 434
Db     955 CTTTTGGAATCAATTCACCCAGCTACACCATGGGCAAAAATGAGCTAAATTGGAAACGGTGAT 896

Qy     435 GTAGCCCTTTGAAAAGTCTTAAAGGGTATCAAGGCATTATATATAGCCATCACCCATCATCAA 494
Db     895 CTAGCCCGTTAAACTCACGCAAATGCACGAAGAACTATCTCTTAATGAATACCGAGCAGCCT 836

Qy     495 TGCCTAAGGCAAGTGCGAGATAAGGCCATTGAGTAGTAGGTTATTGCTTTGCCCTTTGATTTTTAATT 554
Db     835 GTGCTAATGCGATGGCGCCGCCAGTGCGATTCATTATAGGTATTGACGGTCCAAACTAATTTCATCT 776

Qy     555 GATGGGATTGGGTAAAAGTTTTGTTTTTTGATGATATAAATGCCATACCACTACATCATCA 614
Db     775 CATCACAAGCCAAACACCACCCGTTCGTCTTTCGCCCATTCAAATAATGAACCGCCCCTTTTTCGGTTT 716

Qy     615 ATAAAATAAAAAATCTCGCCGTTTGGGTGGCTTAAGTCGTAATTGTTGTTGATAATGTGTGAT 674
Db     715 TTAACCAATAAATCAGCATTGTCTTCCGC-------------------------------- 689

Qy     675 TTGGGGCAACGCTTGTCAGTGGTGGTCAAGCAAGCAATTGCTTAGGCATCGTCGTTGGT 734
Db     688 ------CAAAGCTTACTCCATTTTTTAAATCCGTCCATTGCGGGTTAGTGCATCTTCCGCAT 635

Qy     735 AGATGATTGGGCGGGTTTGGCAATTTTCAAAATAAACACGAAATTTTTGTGTCCAAATAATCCACCA 794
Db     634 TAATCACCGCAGTTTGGGCATGATAGATGATTTTTAATTTTTGCTTCGGGCGGATAGTCTTCTA 575

Qy     795 TGCCATCGTGCCGATCAAGATGATCACGGAGAGATTTAAAAATTGTTGCCCCTTGAGCTG 854
Db     574 AGTCCACATAGCGATTCATCGTGATCTTCACTAATATTTTACACGGTTGCGAAGTGGCTT 515

Qy     855 CCAAAATTGAGATATGCTCAAGCTGAAAAACTTGACGCACCACCAAAAATCCATAT 914
Db     514 TTAAGCTGTAATGTCGTTTCTAGTTGGAAAACTGGTAATTCCAATACATACATATC---AT 458

Qy     915 TTTGGATCGTTAATAATTCAAGTGCAAGGCGGTGCCAATAATTACCGTCCGCGGACACCAACAATCA 974
Db     457 GTGCTTGTTCTTAAATAACGACAGTGGCAATACCAATAATTACCGGCCAATCGGCAACCTTCA 398

Qy     975 TGCCTGCCATGTTTTTGCCATCTCGCCTACTATATGTCGTAACACTGCTTTTGCATTTGAGC 1034
Db     397 AGCCCCGCGCCGCCATTTACCATGTGCGCAACAAGTGATGTGCTTTTTCCCATTTGAGC 338

Qy    1035 CTGTGATGGCGGATGATGATTGGTGTTGGTGTTGCTTGAGTTAGATCTCGTGTATTGAGTGTAT 1094
Db     337 CCGTAATCCGCAATAAATCGGCGTTTTTCGCTTC------------------------------AC 305

Qy    1095 CAATGAACAAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTTTGGCGGCGGACAAATAC 1154
Db     304 GGCAGAACAGCTCAATAATCGGCCAATCACTCAAATACCGCAAGTGCAGTCTGAATTT 245

Qy    1155 TCGGCGTTCTTGGGGTTAATGCCCGGGCTCAATGATGATCTGGCTGAGGTTTAAGAGTT 1214
Db     244 CTGGGTGTTTTTGACCGCTAATCCTGACTAAATAAATCAGATCACTCTCGAGTAACCATT 185

Qy    1215 CACGTGTCCAAGCTGCCAAAATGACAAATGAAGCGCCTGCGGGGGCAGTTTATCTGCCAAGGTAG 1274
Db     184 GTTGGGTTTAAACTGCCAGTATGCAACGGTACATTTTTGCCAATTGTTTCCGGCCCCAGCAG 125

Qy    1275 GATTGGGATTACCATCGATGATGTAACTTTGTGACCGTGATTGGCCAGTTAAGGTTAAATCAAACAG 1334
Db     124 GCTGCGGTTCTCGTATCAAATCACACGCACATCCGCTTAGGCAGGAAAAATCAAACAC 65

Qy    1335 CTGACAGCTTCAGAGCTGCCCAAAACCGACGACGGCATATTTTTT 1377
           || ||| | | |||| | ||   || | | | | || ||||||
Db      64 AAGATAGCCCCTGTTTTCCCAAACCAAATCACGGTAACTTTTTT 22
```

```
Search completed: May 12, 2006, 13:26:10
Job time : 1267 secs
```

OM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 13:05:23 ; Search time 2191 Seconds

        GenCore version 5.1.8
        Copyright (c) 1993 - 2006 Biocceleration Ltd.

Title: US-10-672-787-35_copy_11357_12736
Perfect Score: 1380
Sequence: 1 taaagcgtgaaccatgttaa..........acgacggcatatttctttga 1380
Scoring table: IDENTITY_NUC
          Gapop 10.0, Gapext 1.0

Searched: 9312410 seqs, 2039259788 residues
        (without alignments)
        2568.853 Million cell updates/sec

Total number of hits satisfying chosen parameters: 18624820

Post-processing: Minimum Match 0%
              Maximum Match 100%
              Listing first 45 summaries
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :   Published Applications NA_New:*
          1: /SIDS5/ptcdata/2/pubpna/US08_NEW_PUB.seq1:*
          2: /SIDS5/ptcdata/2/pubpna/US08_NEW_PUB.seq1:*
          3: /SIDS5/ptcdata/2/pubpna/US07_NEW_PUB.seq1:*
          4: /SIDS5/ptcdata/2/pubpna/US06_NEW_PUB.seq1:*
          5: /SIDS5/ptcdata/2/pubpna/US08_NEW_PUB.seq:*
          6: /SIDS5/ptcdata/2/pubpna/PCT_NEW_PUB.seq:*
          7: /SIDS5/ptcdata/2/pubpna/US09_NEW_PUB.seq1:*
          8: /SIDS5/ptcdata/2/pubpna/US10_NEW_PUB.seq2:*
          9: /SIDS5/ptcdata/2/pubpna/US09_NEW_PUB.seq2:*
         10: /SIDS5/ptcdata/2/pubpna/US10_NEW_PUB.seq1:*
         11: /SIDS5/ptcdata/2/pubpna/US10_NEW_PUB.seq3:*
         12: /SIDS5/ptcdata/2/pubpna/US10_NEW_PUB.seq4:*
         13: /SIDS5/ptcdata/2/pubpna/US08_NEW_PUB.seq2:*
         14: /SIDS5/ptcdata/2/pubpna/US11_NEW_PUB.seq2:*
         15: /SIDS5/ptcdata/2/pubpna/US11_NEW_PUB.seq2:*
         16: /SIDS5/ptcdata/2/pubpna/US11_NEW_PUB.seq3:*
         17: /SIDS5/ptcdata/2/pubpna/US11_NEW_PUB.seq4:*
         18: /SIDS5/ptcdata/2/pubpna/US11_NEW_PUB.seq5:*
         19: /SIDS5/ptcdata/2/pubpna/US60_NEW_PUB.seq:*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	160.4	11.6	2925	17	US-11-194-246-220	Sequence 220, App
c 2	132.8	9.6	2850	17	US-11-194-246-205	Sequence 205, App
c 3	132.8	9.6	5640	17	US-11-194-246-225	Sequence 225, App
c 4	78.4	5.7	1323	17	US-11-098-686-10002	Sequence 10002, A
c 5	78.4	5.7	1457619	17	US-11-098-686-8739	Sequence 8739, Ap
c 6	64.6	4.7	1335	10	US-10-467-657-7645	Sequence 7645, Ap
c 7	48.2	3.5	1188	10	US-10-793-626-655	Sequence 655, App
c 8	48.2	3.5	987	18	US-10-793-626-3574	Sequence 3574, Ap
c 9	43.4	3.1	987	18	US-11-079-463-1328	Sequence 1328, Ap
c 10	42.4	3.1	200	17	US-11-098-686-8219	Sequence 8219, Ap
c 11	42.4	3.1	2944528	18	US-11-098-686-8219	Sequence 1, Appli
c 12	42.2	3.0	4010	17	US-11-045-004-1	Sequence 1, Appli
c 13	41	2.9	1362	17	US-10-793-626-3540	Sequence 3540, Ap
c 14	39.4	2.9	3444	10	US-10-793-626-4356	Sequence 4356, Ap
c 15	38.8	2.8	3500	10	US-10-750-185-47750	Sequence 47750, A
c 16	38.8	2.8	3500	10	US-10-750-623-47750	Sequence 47750, A
c 17	38	2.8	432	7	US-09-925-065A-567087	Sequence 567087,
c 18	37.6	2.7	520	7	US-09-925-065A-194762	Sequence 194762,
c 19	37.6	2.7	553	12	US-10-301-480-284002	Sequence 284002,
c 20	37.6	2.7	553	12	US-10-301-480-897411	Sequence 897411,
c 21	37.6	2.7	618	7	US-09-925-065A-916161	Sequence 916161,
c 22	37.6	2.7	620	7	US-09-925-065A-915786	Sequence 915786,
c 23	37.6	2.7	621	7	US-09-925-065A-907030	Sequence 907030, A
c 24	37	2.7	621	10	US-10-750-623-61390	Sequence 61390, A
c 25	37	2.7	877	10	US-10-750-185-61390	Sequence 61390, A
c 26	36.8	2.7	877	10	US-10-240-708-87	Sequence 87, Appl
c 27	36.8	2.7	180531	18	US-10-750-623-61390	Sequence 57, Appl
c 28	36.6	2.7	557	7	US-09-925-065A-637762	Sequence 637762,
c 29	36.6	2.7	180531	18	US-11-114-798-57	Sequence 57, Appl
c 30	36.4	2.6	601	7	US-09-925-065A-517533	Sequence 517533, A
c 31	36.4	2.6	1267	10	US-10-750-623-45019	Sequence 45019, A
c 32	36.4	2.6	1267	10	US-09-925-065A-665713	Sequence 45019, A
c 33	36.2	2.6	535	7	US-10-750-185-45019	Sequence 665713, A
c 34	36.2	2.6	631	11	US-10-301-480-50410	Sequence 50410, A
c 35	36.2	2.6	631	12	US-10-301-480-663819	Sequence 663819, A
c 36	36.2	2.6	2694	10	US-10-750-623-32432	Sequence 32432, A
c 37	36.2	2.6	2694	10	US-10-750-623-32432	Sequence 32432, A
c 38	36	2.6	599	7	US-09-925-065A-645961	Sequence 645961, A
c 39	36	2.6	2337	10	US-10-750-185-26051	Sequence 26051, A
c 40	36	2.6	2337	10	US-10-750-623-26051	Sequence 26051, A
c 41	35.8	2.6	614	12	US-10-301-480-219545	Sequence 219545, A
c 42	35.8	2.6	614	12	US-10-301-480-832954	Sequence 832954, A
c 43	35.8	2.6	621	7	US-09-925-065A-121219	Sequence 121219, A
c 44	35.8	2.6	1241	11	US-10-301-480-40716	Sequence 40716, A
c 45	35.8	2.6	1241	11	US-10-301-480-99957	Sequence 99957, A

ALIGNMENTS

RESULT 1
US-11-194-246-220/c
; Sequence 220, Application US/11194246
; Publication No. US2005027208QA1
; GENERAL INFORMATION:
; APPLICANT: Mott, John
; APPLICANT: Trepod, Catherine
; APPLICANT: Arvidson, Staffan
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND ME
; TITLE OF INVENTION: USE
; FILE REFERENCE: 00592.US1 (M&R 268.05920101)
; CURRENT APPLICATION NUMBER: US/11/194,246
; CURRENT FILING DATE: 2005-08-01
; PRIOR APPLICATION NUMBER: US/10/274,586
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/345,438
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 621
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 220
; LENGTH: 2925
; TYPE: DNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Sequence of the murD coding sequence and flanking regions.
US-11-194-246-220

Query Match   11.6%;   Score 160.4;   DB 17;   Length 2925;
Best Local Similarity   49.6%;
Matches   683;   Conservative   0;   Mismatches   606;   Indels   89;   Gaps   7;

```
Qy      1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGGCTCATTATATAATTTTAAATTGATC  60
          ||||    ||||   |  || |||   || |  |||  |  | | |||   |||||||||||
Db   2240 TAAACATTGAGCTAAAATGCGTAAAATTCTTCGCGCGCCGCTTTTCAAAAGAAGCAAACTGATC 2181

Qy     61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA  120
          ||||||||||||||||||||||||||||||||||||||||||| | ||||   |||  ||
Db   2240 ...
```

```
RESULT 2
US-11-194-246-205/c

; Sequence 205, Application US/11194246
; Publication No. US20050272089A1
; GENERAL INFORMATION:
; APPLICANT: Mott, John
; APPLICANT: Trepod, Catherine
; APPLICANT: Arvidson, Staffan
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND ME
; FILE REFERENCE: 00592.US1 (M&R 268.05920101)
; CURRENT APPLICATION NUMBER: US/11/194,246
; CURRENT FILING DATE: 2005-08-01
; PRIOR APPLICATION NUMBER: US/10/274,586
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/345,438
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 621
; SOFTWARE: Patentin version 3.0
;
; SEQ ID NO 205
; LENGTH: 2850
; TYPE: DNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Sequence of the mraY coding sequence and flanking regions.
US-11-194-246-205

  Query Match        9.6%;  Score 132.8;  DB 17;  Length 2850;
  Best Local Similarity  54.0%;  Pred. No. 5.1e-25;
  Matches  337;  Conservative  0;  Mismatches 257;  Indels  30;  Gaps  2;
```

```
Db  2180 GAGACTTGCAAGCAGGCGACGACAATAATACCA-----------TATCTCGGCTTTGCAA 2133

Qy   121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGAGTGGCATCACCTTT        180
Db  2132 TGTTGGGCCGTAAAAATTCTATGCGTTGTTCCATTGTGTATCGAACAA------ATAACTTTG 2079

Qy   181 GAGTAAAATCTTGCTCAATTATCGGGGGGCATCGTTTGACCGGTGATATTCTCCACCCA    240
Db  2078 CGATGAAAAATTTGCAAGCAGGCGACCACCATCTCGACCAAAACAATAACAAATAAAGTGTGG 2019

Qy   241 CTGCATAACATATGCTGACAGCTCACTATAAATCTTGACCTTTACCCAAGCCCACCCAAAAAT 300
Db  2018 TTGATTAATTAATTCAGCTAATTCTGAAAAAAATCAGCCCCTTTTTCCGTTCTCCGGCAA    1959

Qy   301 CAATGCCAGTGATGATATGCCATAAAACCGTGCCAAGCCCATCAATTGCTGCAAATGGTTGA  360
Db  1958 CAAATGCAATTTACCCTCAATATA-------AAGCCAGCCAATGCACCAACTGTACT       1908

Qy   361 GCCGGATGTTTGTACCTTTTGAAGTCAATCAATAAAAATATGCCTTGCTGTATCAATATCATCAATATA 420
Db  1907 CCCCACATTGTTGCTTTAGAGTCATTATTAATCCAACGAAATGCCATTAGCTTTGATGCACTAA 1848

Qy   421 TTCACAGCCGATGCGGTAGCCCCTTTTGAAAGTTCTTAAGGGTATCAAGCACATGGCATCCATCATGGG 480
Db  1847 TTGAAAACGGATGATCTAACCCTTTGAAAAATGACGAAGTGCGGGTACCGAAATCGAAATCTAAATT 1788

Qy   481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCCAGATAAAGGCATTGAGTAGGTTATTAGCGTTGCC  540
Db  1787 AATACCTATAGCTGTGTCCAAATGCTTAGTTGCTCGCCAAAATGTTCATTATTATTATGCGACC 1728

Qy   541 TTTGATTTTTAAATTGGATTGGTAAAAGTTTGTTTTTTGATGATATAATGCCATACC        600
Db  1727 AACCAATGTAGCTTCTTCACAAGGTAAAAATCACTTCATCTTTACCATTAAAATATGCTTT 1668

Qy   601 ATCACCATCATCAAAATAAAAATAAAACTCGC--CGTTTGGTGTGGCTAAGTAAGCTATTTGTTG  659
Db  1667 GCCATTTTCAGTTTTTAGCCAATAATCCCGCACTATTTTCCGCAAAAGAAACGGTATGTTTT    1608

Qy   660 TGAATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCAATTGCTTGGCTAAAT    719
Db  1607 CGCTTGATTTTTCGTTTTTCCCCAAAAGTCAGCC------------                 1576

Qy   720 TGGCATCGTTCTTGGGTAGATGATGGCGGTTTGGCAATTTTCAAAAAATACCGCAATTTTGTTG   779
Db  1575 ---TATCTTCATTGTTCAACACATTGCATTATGATTAATGCCTTAAATGTCGTAACAGTGC     1519

Qy   780 CCAAATAATCCACCATGCCATCGTCCGATCAAGATGATCAGCGGAGAGATTTAAAATTG        839
Db  1518 GGGCGATAATCTTCTAAAATCCATAATAGCGGATTCCAATATGACTTCAGTTCAAGACAG 1459

Qy   840 TTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCA     899
Db  1458 TCGGTGCCCCAGCTTTTTAAGGCTAATAAGTTGTCTCAAGCTGAAAAACTAGAAAAGCTCTAGTA 1399

Qy   900 CCACCAAATCCATATTTGGATCGTTAATATAATTCAAGAGTGCCAAGGGTGCCAATATTTACCGC   959
Db  1398 CATAAAGTTCACAAATCTTCATTC----AACAATGACAAAGCGAATCCCAATAATTCCCG     1342

Qy   960 CGACACCAACAATCATGCCTGCAATTTGCCATCTCGCGCTACTATTTAGCAATTTGTAACAGTGC 1019
Db  1341 CCATACCAACTTTCACACAGCTTTCGCCATTTCATAAAACTAAAGATGTTAGGGATCTC   1282

Qy  1020 TTTTGGCATTTGAGCTGCCTGATGATGTAGAGTCAGTTAGAGATCTC                 1079
Db  1281 TTTTACCATTTGAACCTGTAATCCCCACAATGGCCTTGAGTGTTGTTCGCTTGAGATCTG 1230

Qy  1080 GTGCTTGAGTGTATCAATGAACAAATGACGCTGATCACAGGAATGCCTTGAGCTT        1139
Db  1229 -----------GAATAAATTCAATATACGGCGATTACTTCCAATCATATGCC           1189

Qy  1140 TGGCGGGCGACAAATACTCGCGGCGGTTCTTGGGTTAAATGCGGGGGCTGATGATGAATCTTCGG  1199
Db  1188 GTCGGGTTTGAAATTCGGTTCTTGGTGTTTTACCGGCAAGCCCTGGGCTTAATAACAATCATATGC 1129

Qy  1200 CTGAGCTTAAGAGTTTCACTGTCCAAGCTGCCAAATGACAATGAAGCCCTGTGGGCAGTT   1259
Db  1128 TTTCAAGTAACCATTCCTGATTTAAAACTACCAGTATGAAGAGGAATTTTGGAGAAGTT    1069

Qy  1260 TATCTGCCAAGGTAGGATTGGGGATTACCATCCGATGAGTGCTAACTTTGTGACCGTGATTGA 1319
Db  1068 TATCAATACCAGTAGGATTTTTTCGAGTATCAAATCACACGAATTAATAGCCTTGTTCGTTGGGATA 1009

Qy  1320 CCAAAAAATTCACAGCTGTCAGAGTCCCAAAACCGACGACGGGCGGCCATATTTTT    1377
Db  1008 AGAGATAATCCACACACAAGAAAGACCCTGTTTTGCCAAGCCCGATGATGAGTAATATATTTTT 951
```

```
Qy   754 ATTTTCAAAATAATACCGCAATTTTTGTTGTCCAAAATAATAATCCACACATGCCATCGTCCGATCAAG 813
Db  2839 ATTATGATAAAATGGCGTAAATTTTTGCTTGGGCCGGATAAATCCCATAATAGCGGATCCAT     2780

Qy   814 ATGATCAGCGGAGAGAGATTTAAAATTGTTGCCCCTTGAGCTGCCAAAATTTGAGATATGCTC 873
Db  2779 ATGATCTTCAGTCACGTTCAAGACAGTCGGTGCGCCCCAGCTTTTTAAGCTAATAAGTTGTCTC 2720

Qy   874 AAGCTGAAAACTTGACAGCTCAAGCACCACCAAATCCATATTTTGGATCGTTAATATAAATTC    933
Db  2719 AAGCTGAAAAACTAGAAAAGCTCTAGTACATAAAGTTCACAAATCTTCATTC---AACAATGA 2663

Qy   934 AAGTGCCAGGGGGTGCCAATATTACCCGCGACACCAACAATCATGCCTGCAATTTGCCAT     993
Db  2662 CAAAGCGGGAATCCCAATAATTCCCGCCATACCAACTTTCACACAGCTTTCGCCAT          2603

Qy   994 CTCGGCCTACTATTTAGCAATGTCGTAACAGTGCTTTTGGCATTTGAGCCTGTGATGGCGATGATGG 1053
Db  2602 TTCATAAAACTAAAGATGTTAGGGATCTCTTTTTACCATTTGAACCTGTAATCCCCACAATGG 2543

Qy  1054 TGTGTTTGTTGCTTGAGTTTAGATCTCGTGCTTGAGTGTTGTTCGCTTGAGATCATC       1113
Db  2542 CTTTGTCGCTTGCGCG-----------GCAGAATAATTCAATATATC                2510

Qy  1114 GCTGATCACAGGAATGCCTTGAGCTTTGGCGGCGACAATACTCGCGGCGGTTTCGTCTTTGGGTTTAAT 1173
Db  2509 GCCGATTACTTCCACTCCCCGCCCTTTAAGTGCGGTTTGAATTCGGTGTTTTACCGCAAG    2450
```

RESULT 3
US-11-194-246-225/c
; Sequence 225, Application US20050272089A1
; Publication No. US20050272089A1
; GENERAL INFORMATION:
; APPLICANT: Mott, John
; APPLICANT: Trepod, Catherine
; APPLICANT: Arvidson, Staffan
; TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND MET
; TITLE OF INVENTION: USE
; FILE REFERENCE: 0592.US1 (M&R 268.05920101)
; CURRENT APPLICATION NUMBER: US/11/194,246
; CURRENT FILING DATE: 2005-08-01
; PRIOR APPLICATION NUMBER: US/10/274,586
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/345,438
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 621
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 225
; LENGTH: 5640
; TYPE: DNA
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: Sequence of the murE, murF and mraY coding sequences and flanking
; OTHER INFORMATION: regions
US-11-194-246-225

Query Match   9.6%;   Score 132.8; DB 17; Length 5640;
Best Local Similarity   54.0%;
Matches 337; Conservative   0; Mismatches 257; Indels 30; Gaps   2;

```
Qy   1174 GCCGGGGGCTGATGATGATCTCGGCGTCGCTGAGCTTAAGAGTTCACTGTCCCAA 1233
          |||||      ||| |     |||||||      ||||  |||| |||||||
Db   2449 CCCTGGGCGCTAATAACAATCATATCCGCTTCAAGTAACCATTCCCTGATTAAACTACCAGT 2390

Qy   1234 ATGACAATGAACGCCTGTGGCGCAGTTTATCTCGCCAAGGTAGGATTGGGGATTACCATCGAT 1293
          ||||  |||    |  |  |||||||||||||||||||||||||| || ||||||||||||||
Db   2389 ATGAAGAGGGATATTTTTGAGAGAAGTTATCAATACCAGTAGGAGTTTTTCGAGTATCAAT 2330

Qy   1294 GATGCTAACTTTGTGACCGTGATTGACCAAAAAATTCACAGCTGACAGTCCAGAGCTGCC 1353
          ||      ||   |     |||| ||||||||    ||||||||||||||||
Db   2329 CACACGAATATTAGCCTGTTGGGATAAGAGATAATCCACACAAGAAAGACCTGTTTGCC 2270

Qy   1354 CAAACCGACGACGGCATATTTTTT 1377
          ||||||||||||||||||||||||
Db   2269 AAGCCCGACGATGAGTAATATTTTTT 2246
```

RESULT 4
US-11-098-686-10002/c
; Sequence 10002, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10002
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10002

Query Match   5.7%;   Score 78.4;   DB 17; Length 1323;
Best Local Similarity   54.2%;
Matches 182; Conservative   0; Mismatches 151; Indels   3; Gaps   1;

```
Qy   754 ATTTTCAAAAAATACGCAAATTTTTGTGCCAAATAATCCACACATGCCATCGTCGCCGATCAAG 813
         |||||          ||||||||  |||||||||||||||||||    ||||| ||||||| ||
Db   666 ATTTTTACATAATATTCAACTTTGCAGAAAATATTCTATGGTAATCCAA 607

Qy   814 ATGATCAGCGGAGAGATTTAAAATTGTTGCCCCTTGAGCTGCCAAATTTGAGATATGCTC 873
         ||||||                   || |||        | |   |  |  ||||||||||
Db   606 ATGATTTATAGAAAATATTAGAACACATATAGCTATATCGAGGACGAAAATAGTGAACATGTCTG 547

Qy   874 AAGCTGAAAACTTGACAGCTCAAGACACCACCAAAATCCATATTTTTGGATCGTTAATAAATTC 933
         ||||||||||||||||    |  |||||||  |||||||||||||||  |||| ||   |  |
Db   546 AAGTTGAAAACTAGAAAGTTCACAAATCTCATTC---AACAATGA 487

Qy   934 AAGTGCAGGCGGCTGCCAATATTACCGCGACCAACAACAATCATGCCTGCCAT 993
         |  | | |||| || ||||||||  ||||   |||| ||  |||| ||||||
Db   486 TTCTGATAGTGGAGTTCCAATATTACCACAACAACACTTATACCTTGTTCTACCAA 427

Qy   991 CATCTCGCCTACTAAGTCGTAACAGTGCTTTTGGCATTTGACGGCGGATGAT 1050
         || |  |||| |   ||| | || ||  |||||| ||||||| |||||||||
Db   426 CATTGCTGCACATAAACTTGCTATAGTTGTTTTCCCATTGGTCCCTGTTTATGGCAATGAC 367

Qy  1051 TGGTGTTGTTTGTTGCTTGAGTTAGATCTCGTTGCTTT 1086
         ||||| ||||||| ||||||||| ||  | || ||| |
Db   366 TTTTTCATTAGACAACTGATACCAATGCTAGTTCTGT 331
```

Db   366 TTTTTCATTAGACAACTGATACCAATGCTAGTTCTGT 331

RESULT 5
US-11-098-686-8739/c
Sequence 8739, Application US/11098686
Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8739
; LENGTH: 1457619
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-8739

Query Match   5.7%; Score 78.4; DB 17; Length 1457619;
Best Local Similarity 54.2%; Pred. No. 1.6e-09;
Matches 182; Conservative 0; Mismatches 151; Indels 3; Gaps 1;

Qy   754 ATTTTCAAAAAATACGCAATTTTTGTGCCAAATAAATCCACCACCATGCCATCGTGCCGATCAAG 813
              || ||| | ||| ||| | ||| | | | ||| | ||| || || | | |||| |
Db   1356224 ATTTTTACATATATTCAACTTTGCAGAAATATATTCATCCATATTTCTATGGTAATCCAA 1356165

Qy   814 ATGATCAGCGGAGAGAGATTTAAAATTGTTCGCCCTTGAGCTGCGCCAAAATTTGAGATATGCTC 873
              ||| | | || || |||| || | | | |||| | || | |||| || ||| | |
Db   1356164 ATGATTTATAGAAAATAATTAGAAACATATAGGCTATATCAGGACGAAATAAGTGAACANGTCTG 1356105

Qy   874 AAGCTGAAAACTTGACAGCGTCAAGCACCACCAAAATCCATATTTTTGGATCGTTAATA---A 930
              || ||| | || | || || | || |||||| || | ||| | | | |
Db   1356104 AAGTTGAAAACTTCTAAAAACAAACAACTACTTCTTTCTTTTTTACACAAAAACATA 1356045

Qy   931 TTCAAGTGCAAGGCGGTGCCAAATATTACCGCGACAACAACACTTACTTCTTCTGGTCCCAGTTTTGC 990
              ||| ||| | | || || ||| |||||| | | || || | ||| | || || | ||| ||
Db   1356044 TTCTGATAGTTGGAGTTCCAATAATTACCACCAACAAAAACACTTTATTACCTTCGTTCTTCTACCAA 1355985

Qy   991 CATCTCGCCTACTAATGTCGTAACAGTGCTTTTGGCATTTGAGCGTGTCGATGGGCCATGAT 1050
              || | | || | || | || | | ||| | ||| | | ||| | ||||| | |
Db   1355984 CATTGCTCGCACATAAACTTGCTATAGTGTTGTTTTCCCATTGGTCCCTGTTATGGGCAAATGAC 1355925

Qy   1051 TGGTGTGTTTGTTGCTTGAGTTAGATCTCGTGCTTT 1086
              | || |||| || |||| | ||| | | | |||
Db   1355924 TTTTTCATTAGACAAACTGATACCATTGCTAGTTCTGT 1355889

RESULT 6
US-10-467-657-7645/c
Sequence 7645, Application US/10467657
Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 7645
; LENGTH: 1335
; TYPE: DNA

Right column:

; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-7645

Query Match   4.7%; Score 64.6; DB 10; Length 1335;
Best Local Similarity 49.9%; Pred. No. 8.3e-07;
Matches 203; Conservative 0; Mismatches 189; Indels 15; Gaps 1;

Qy   185 AAATCTTGCTCAATATATCGGGGATGGGCCATCTTGACCGGATTAAAATACACACTATCCACCACTGC 244
Db   1157 AAATCGCGGCGGGATTTGCGCGCGCCATCGCACGCGCCGATCAGGAAACACGCCTTTTTGCCTTATCT 1098

Qy   245 ATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCCAAAATCAAT 304
Db   1097 TTCAACGCGTGCGGCGCGAGGGCGGTGAAGGTCCTGCCCCTTTGCCCATGCGGCCCAAAATCACG 1038

Qy   305 GCCAGTGATGATATGCCATAAAACCCTGCCAAACCCATCAATTGCTGCCAATGGTTGAGCCG 364
Db   1037 AAGAGCG------------GGTTTGCAAACCGGCAATCGCGCGGGGGGGGGGTCGGCCCG 993

Qy   365 ATGTTTGTACCTTTGAAATCATTAAAAATATGCCTTGCTCTATCAAATATCAATAATAATATTCA 424
Db   992 ACATTCGTGCCTTTGCTGTGCGTGATGAAATAACCACGCCGTTTTCTCGCGCGGATTTTTTCC 933

Qy   425 CAGCGATGCGCGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATACGGCATCCATGGGGAAGA 484
Db   932 ACACGGTGCGCGACAAGCCTTGGAAGGTTTTGACGGTGTTCCAGCAATGCTTGGCGCGCGCAAA 873

Qy   485 TTGGCAGCTGTGCCTAAGGCAAGTGCGAGATAAGGCATTGAGTAGGTTATGCTTGTGCTTTG 544
Db   872 CCGACGGCCTCGCACAAAGCCACGCGGCCATAACGTTGGCGCGGTTGTGCAAACCTTGC 813

Qy   545 ATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATA 591
Db   812 AGCGGGATGTCTTGCGTAGAAATCAAAATCTTCATTGCCTTGTTTTTPA 766

RESULT 7
US-10-793-626-655/c
Sequence 655, Application US/10793626
Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU340US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 655
; LENGTH: 1188
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-655

Query Match   3.5%; Score 48.2; DB 10; Length 1188;
Best Local Similarity 53.4%; Pred. No. 0.02;
Matches 101; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Qy   366 TGTTTGTACCTTTTGAATACATTAAAATATGCCTTGCTATCAAATCATCAAAATATATTCAC 425
Db   832 TATTAGTTGCTTTGAATCATTATAATATTTATTTGTGGGATTTGTTACCAATATACTGAA 773

Qy   426 AGCGGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGCCATCCATGGGGAAGAT 485
Db   772 GTCTATGATCAAATACGGAGAAAAGTAACAAGACTATCTACAAATAGCTTTGACTGGCACTC 713

Qy   486 TGGCAGCTGTGCCTAAGGCAAGTGCGAGATAAGGCATTGAGTAGGTTATGCTTGTGCCTTTGA 545

```
Db    712 CAGCAATGATTGATGCTAGAACAGCTGCTAAAAATATTTCCAGGTTATGTTCTCCTGGTA 653
              |||
Qy    546 TTTTTAATT 554

Db    652 GCACTAAGT 644
```

RESULT 8
US-10-793-626-3574
; Sequence 3574, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 3574
; LENGTH: 3334
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3574

Query Match   3.5%; Score 48.2; DB 10; Length 3334;
Best Local Similarity 53.4%; Pred. No. 0.029;
Matches 101; Conservative   0; Mismatches 88; Indels   0; Gaps   0;

```
Qy    366 TGTTTGTACCTTTTGAATCATTAAAAATATGCCTTGCTATCAATATCATCAATATATTCAC 425
              |||||||||||||||||| ||||||||||  |||||||||||  ||| |||||
Db    299 TATTAGTTGCTTTTGAATCATTATAATATTATTGTGCCAATATACTTGTACCAATATACTGAA 358

Qy    426 AGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGTATCAAGCGATGGCATCCATGGGAAGAT 485
              |||  ||||||||||| ||||||||||||
Db    359 GTCTATGATCAATACCGGAAAAAGTAACAAGACTATCTACAATAGCTTTGACTGGCACTC 418

Qy    486 TGGCAGCTGTGCCTAAGGCAAGTGCGAGATAAGGCATTGACTGAGTTATGCTTGCCTTTGA 545
              ||||||||||||||||||||||
Db    419 CAGCAATGATTGATGCTAGAACAGCTGCTAAAAATATTTCCAGGTTATGTTCTCCTGGTA 478

Qy    546 TTTTTAATT 554
              |||||
Db    479 GCACTAAGT 487
```

RESULT 9
US-11-079-463-1328/c
; Sequence 1328, Application US/11079463
; Publication No. US20060073161A1
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO BACTEROIDES FRA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: PATH00-03DIV2
; CURRENT APPLICATION NUMBER: US/11/079,463
; CURRENT FILING DATE: 2005-03-14
; PRIOR APPLICATION NUMBER: US 60/128,705
; PRIOR FILING DATE: 1999-04-09
; PRIOR APPLICATION NUMBER: US 09/540,209
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 10444
; SEQ ID NO 1328
; LENGTH: 987
; TYPE: DNA
; ORGANISM: B.fragilis
; FEATURE:
; NAME/KEY: unsure

; LOCATION: (928)..(942)
; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkn
US-11-079-463-1328

Query Match   3.1%; Score 43.4; DB 18; Length 987;
Best Local Similarity 48.8%; Pred. No. 0.37;
Matches 147; Conservative   0; Mismatches 151; Indels   3; Gaps   1;

```
Qy    754 ATTTTCAAAAATACGCAAATTTTTGTGGCAAATAAATCCACCATGCCA---TCGTGCCGATC 810
              |||||||| |||||||| ||||     |||||||||||||||||||||     ||||| ||||
Db    759 ATTCTCGGTAATACGAAACTTTGCATTAATAATAGTTCTGCATATACAATGGTCGTCGTACCGGTC 700

Qy    811 AAGATAGATCAGCGGAGAGATTTAAAAATTGTTGTCCCCTTGAGCGTGCCAAAATTTGAGATATG 870
              |||||  |||| ||  ||| |||   ||||||||||||||
Db    699 CAGATGGGTCCGGCGTAATGTTCATCAATACACGACGATATCGGGCACGGGAAGTTATACATGTT 640

Qy    871 CTCAAGCTGAAAACTTGACAGCTCAAGCACCACCACCAAAATCCATATTTTGGATCGTTAATAA 930
              |||||||| ||    ||||  |||||||||||||||||||||||||||||||||||||||||||
Db    639 ATCCAACTGGAATGAACTCAATTCGATTACATTACATTAAATAATCATGTTTCTCTTTCGGCCACTTG 580

Qy    931 TTCAAGTGCAGGCGTGCCAATATTACCGCCGACACCAAACAATCCATGCCTGCAGTTTTGC 990
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    579 CAATGCCAGGCTTTTACCGATGTTTCCGGCAAGTCCCACATTTAGTCCGCGCTTTTAAA 520

Qy    991 CATCTCGCTACTAAGTCGTAAACACTGTCTTTTGGCATTTGAGCGTCTGATGCGATGGAT 1050
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    519 AATGTGATTAGGATAAGCCAGAGGTTGTGGGTCGTCTTTCCGTTCGAGCCGGTGATACAGATCAT 460

Qy    1051 T T 1051
               |
Db    459 T T 459
```

RESULT 10
US-11-098-686-8219/c
; Sequence 8219, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 8219
; LENGTH: 200
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-8219

Query Match   3.3%; Score 42.4; DB 17; Length 200;
Best Local Similarity 54.6%; Pred. No. 0.41;
Matches 107; Conservative   0; Mismatches 86; Indels   3; Gaps   1;

```
Qy    782 AAAATAATCCACCATGCCATCGTGCCGATCCAAGATAGATCAGCGGAGAGATTTAAAATTGTT 841
              ||| ||||||| ||||| ||| |||  ||||||||| || |||||||| |||||  |||||||
Db    196 ATATATTCATCCATCATTTCTATGGTAATCCAAATGATTTATAGAAATATTAGAACATATA 137

Qy    842 GCCCCTTGAGCGTGCCAAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCGACC 901
              ||  ||||||||    |||  ||||  ||||| ||||||||||||||||||  || |||  ||
Db    136 GCTATATCAGGACGAAAATAGTGAACATGTCTGAAGTTGTGAAAAACTTGAAAGTTCTAAAACA 77

Qy    902 ACCAAATCCATATTTTGGATCGTTAATA---ATTCAAGTGCAGGCGTGCCAATATTACCG 958
              ||  |||||||| ||||| || ||||||    |||||||| || ||||||||||||||||
Db    76 ACAACACTTACTTCTTTTTTACACAAAACATATTCTGATAGTGTGGAGTTCCAATATTACCA 17

Qy    959 CCGACACCAACAAATCA 974
```

```
Db       16  CCAACAAAAACACTTA 1
             ||| |||     |||| ||
Qy        3  ..........  1
```

RESULT 11
US-11-045-004-1
Sequence 1, Application US/11045004
Publication No. US20060078901A1
GENERAL INFORMATION:
; APPLICANT: BUCHRIESER, CARMEN
; APPLICANT: FRANGEUL, LIONEL
; APPLICANT: COUVE, ELISABETH
; APPLICANT: RUSNIOK, CHRISTOPHE
; APPLICANT: FSIHI, HAFIDA
; APPLICANT: DEHOUX, PIERRE
; APPLICANT: DUSSURGET, OLIVIER
; APPLICANT: CHETOUANI, FARID
; APPLICANT: NEDJARI, HAFED
; APPLICANT: GLASER, PHILIPPE
; APPLICANT: KUNST, FRANCK
; APPLICANT: COSSART, PASCALE
; APPLICANT: DANIELS, JUSTIN
; APPLICANT: GOEBEL, WERNER
; APPLICANT: KREFT, JURGEN
; APPLICANT: KUHN, MICHAEL
; APPLICANT: NG, EVA
; APPLICANT: VAZQUEZ-BOLAND, ANTONIO
; APPLICANT: DOMINGUEZ-BERNAL, GUSTAVO
; APPLICANT: GARRIDO-GARCIA, PATRICIA
; APPLICANT: TIERREZ-MARTINEZ, ALBERTO
; APPLICANT: AMEND, ALEXANDRA
; APPLICANT: CHAKRABORTY, TRINAD
; APPLICANT: DOMANN, EUGEN
; APPLICANT: HAIN, THORSTEN
; APPLICANT: BERCHE, PATRICK
; APPLICANT: CHARBIT, ALAIN
; APPLICANT: DURANT, LIONEL
; APPLICANT: PEREZ-DIAZ, JOSE-CLAUDIO
; APPLICANT: BAQUERO, FERNANDO
; APPLICANT: GARCIA DEL PORTILLO, FRANCISCO
; APPLICANT: GOMEZ-LOPEZ, NURIA
; APPLICANT: MADUENO, ENCARNA
; APPLICANT: PABLOS, BETRIZ DE
; APPLICANT: WEHLAND, JURGEN
; APPLICANT: KARST, UWE
; APPLICANT: ENTIAN, KARL-DIETER
; APPLICANT: HAUF, JORG
; APPLICANT: ROSE, MATTHIAS
; APPLICANT: VOSS, HARTMUT
; TITLE OF INVENTION: LISTERIA MONOCYTOGENES GENOME, POLYPEPTIDES AND USES
; FILE REFERENCE: 05394.0018-02
; CURRENT APPLICATION NUMBER: US/11/045,004
; CURRENT FILING DATE: 2005-01-28
; PRIOR APPLICATION NUMBER: 10/637,657
; PRIOR FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: 10/257,023
; PRIOR FILING DATE: 2002-10-08
; PRIOR APPLICATION NUMBER: PCT/FR01/01118
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: FR 00/04,629
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 2854
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 1
; LENGTH: 2944528
; TYPE: DNA
; ORGANISM: Listeria monocytogenes

Query Match   3.1%; Score 42.2; DB 18; Length 2944528;
Best Local Similarity   50.2%; Pred. No. 11;
Matches 104; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Qy   334 AAGCCCATCAATTGCTGCAAATGGTTGAGCCGATGTTTGTACCTTTTGAATCATTAAAATA 393
             | | ||||| |||| |||| |||||| |||| | | || |||||| | || |||||
Db 2117634 AAACCCTTTTAATGCACTTTGAGTTGCAAGGATAATTCGTTGCTTTTTGAATCGTTATAGAA 2117693

Qy   394 TGCCTTGCTATCAATATTCATCATCAAATATATTCACAGCGATGCGGTAGCCCTTTGAAAGTCTT 453
             ||||| || | |||| | ||| || || || || | |||| | ||||| |
Db 2117694 TTTACGACCTTGCCATTCTACACACAAAATTGTGTCCGGTGCTCTACACCTTTAAATGTTTC 2117753

Qy   454 AAGGGTATCAAGCATGGCATCCATGGAGAAGATTGGGCAGCTGGTGCCTAAGGCAAGTGCGAGA 513
              || ||| |||| ||||| | || |||| ||| || ||||| || ||||| |||||
Db 2117754 TAATACGTGCATGATTTCTTCGTTTGTAACACCTAAATGTCTTTGCTACTGCTACGGAAAGC 2117813

Qy   514 TAAGGCATTGAGTAGGTTATGCTTGCC 540
             || ||| |||||||| ||| || |||
Db 2117814 AAGTACGTTTTCTAAGTTATGTTCTCC 2117840
```

RESULT 12
US-10-793-626-3540
Sequence 3540, Application US/10793626
Publication No. US20050254784A1
GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3540
; LENGTH: 4010
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3540

Query Match   3.0%; Score 41; DB 10; Length 4010;
Best Local Similarity   48.2%; Pred. No. 2.6;
Matches 147; Conservative 0; Mismatches 155; Indels 3; Gaps 1;

Qy   754 ATTTTCAAAAATACGCAAATTTTTGTGCCAAATAATCCACCATGCCATCGTGCCGGATCAAG 813
             |||| | || || | |||||| ||| |||| ||| || | || | | |||| |||
Db   42 ATTTTTATATATTTGCTTTTTAGCATTTGATAGTTCTCTAAACGTTTTCANGGTAATCCAA 101

Qy   814 ATGATCAGCCGAGAGATTTAAAAATTGTTGCGCCCTTGAGCGTGCCAAAATTTGAGATATGCTC 873
             |||| | || | || ||| | | |||| | |||||| | | || ||||| || |||
Db   102 ATGTGCAGAATAAATATTAGTAATGATAGCGCAATGTGGTTTATATATTCTTCAAATGCCTAA 161

Qy   874 AAGCTGAAAAACTTGACAGCTCAAGCGACCACCACCAAAATCCATATTTGGATCGTTATT---AA 930
             || |||||| | || | |||||| | ||| | | |||| || || |||| || ||
Db   162 TAATTGAAAAGATGATAATTCTGTTATTAAAATACTCATCTGATTTAAACTTCTTGTCAAC 221

Qy   931 TTCAAGTGCAGGCGGTGCCCAATATTACCGCCGACCACCAACAAATCATGCCTGCCATGTTTTGC 990
             | ||| | | |||| | || | |||| || | | | ||| | | || | ||| ||
Db   222 TTTTGAGGCGTACATAACCAATAATTCCCAGAAAAGTCGTCAACACGCTTTTTGGAA 281

Qy   991 CATCTCGCCTACTAAATGTCGTAACAGTGCTTTGCATTTGAGCCTGTGATGGGGGATGAT 1050
             | ||| || ||| | || ||| | ||| || |||||| || || |||| |||||
Db   282 AATATCACCGATTAGTGACAGTAGTTTACCGTTAGTTCCAGTAAACTGCTATGAT 341

Qy   1051 TGGTG 1055
            |||||
Db   342 TGGTG 346
```

RESULT 13
US-11-098-686-10005/c
Sequence 10005, Application US/11098686
Publication No. US20060024696A1

RESULT 15

GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: US/11/098,686
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SEQ ID NO 10005
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10005
; SOFTWARE: FastSEQ for Windows Version 4.0

Query Match             2.9%;  Score 40;  DB 17;  Length 1362;
Best Local Similarity   53.1%;  Pred. No. 3,4;
Matches    85;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy    381 AATCATTAAAATATGCCTTGCTATCAAATATCATCAATATATTCACAGCGATGCGGTAGCC 440
Db    979 AATCATCAATAACACAGTAATTCCATTCTTTCCCCTTTATATTCAAAAGCGTTCCTACCC 920

Qy    441 CTTTGAAAGTCTTAAGGGTATCAACGCATGGCATCCATGGGAAGAAGTTGCAGCTGTGCCTA 500
Db    919 CTTTAAAGTTAGCAAGTCCCCTCAATACAACATTGGGTGAAATACCTACTTCTAATGCTA 860

Qy    501 AGGCAAGTGCGAGATRAGGCATTGAGGTAGGTTATGTGCTTGCC 540
Db    859 CACCTTATTGCAGCTAATGCATTAAGAAATAATTATATGTTTCCC 820

RESULT 14
US-10-793-626-4356/c
Sequence 4356, Application US/10793626
Publication No. US20050255478A1
GENERAL INFORMATION:
; APPLICANT: KIMBERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4356
; LENGTH: 3444
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-4356

Query Match             2.9%;  Score 39.4;  DB 10;  Length 3444;
Best Local Similarity   65.2%;  Pred. No. 6,7;
Matches    58;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy      5 GCGGTGAACCATGTTAACAAACTTACACACTCCGCTCATTATTAATTTTTAAATTGATCCATG 64
Db    378 GCTCGGAATCTATCCGATAAACTTCTCACCAGGTTCTTCAAATGTATGATGATCCAA 319

Qy     65 CTAGCACAAGCCAGGTTGACAAAAGCAGC 93
Db    318 CTTGCACAAGCCTGGTTGATAAATAGAACAAC 290

US-10-750-185-47750
Sequence 47750, Application US/10750185
Publication No. US20050260603A1
GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 47750
; LENGTH: 3500
; TYPE: DNA
; ORGANISM: Bovine          19866880704052
US-10-750-185-47750

Query Match             2.8%;  Score 38.8;  DB 10;  Length 3500;
Best Local Similarity   51.8%;  Pred. No. 9,7;
Matches    91;  Conservative    0;  Mismatches   87;  Indels    0;  Gaps    0;

Qy    364 GATGTTTGTACCTTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATATTC 423
Db   1048 GATGTGACTAACATTTAAAATCACTAGACTTTGAGTAAAGACAAATTATCTTCATAATGTA 1107

Qy    424 ACAGCGATGCGGTAGCCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGGAAG 483
Db   1108 ATAGGTCTGACCTCATCCATTGAAGCCCTTAAGAGAGAAAAATTACTGAGATTCCCTGAGAA 1167

Qy    484 ATTGGCAGCTGTGCCTAAGGCAAGTGCGAGATAAGGCATTGAGGTAGGTTATGCTTGCCT 541
Db   1168 GAAGTGCTCTGACTCTAAATTGCCTTCAGAGTCAGTACTGCTACATTAGCTGCTTCCCT 1225

Search completed: May 12, 2006, 14:02:35
Job time : 2200 secs

THIS PAGE BLANK (USPTO)

| | |
|---|---|
| **From:** | Chan, Christina |
| **Sent:** | Tuesday, May 16, 2006 3:51 PM |
| **To:** | Graser, Jennifer; STIC-Biotech/ChemLib |
| **Subject:** | RE: rush search |

Please rush. Thanks Chris

**Chris Chan**
TC 1600 New Hire Training Coordinator and SPE 1644
(571)-272-0841
Remsen, 3E89

-----Original Message-----
**From:**     Graser, Jennifer
**Sent:**     Tuesday, May 16, 2006 2:43 PM
**To:**       Chan, Christina
**Subject:**  rush search

Hi Christina,

Will you please authorize the following rush search a special case which is due?

Thanks,
Jennifer
--------------
STIC:
Please search oligonucleotides of at least 22 nucleotides in length from 11357 to 12736
of SEQ ID NO: 35 from application 10/672787 in pending and commercial databases.
Thanks,
Jennifer Graser
Primary Examiner
Art Unit 1645
REM 3B09
2-0858

| ******************** | ******************** | ******************** |
|---|---|---|
| Searcher:_____ | **Type of Search** | Vendors and cost where applicable |
| Searcher Phone:_____ | NA#_____ AA#:_____ | STN:_____ |
| Date Searcher Picked up:_____ | S/L:_____ Oligomer:_____ | DIALOG:_____ |
| Date completed:_____ | Encode/Transl:_____ | QUESTEL/ORBIT:_____ |
| Searcher Prep Time:_____ | Structure #:_____Text:____ | LEXIS/NEXIS:_____ |
| Online Time:_____ | Inventor:_____ Litigation:___ | SEQUENCE SYSTEM:_____ |
| | | WWW/Internet:_____ |
| | | Other (Specify):_____ |

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: May 18, 2006, 08:45:02 ; Search time 947 Seconds
(without alignments)
10160.193 Million cell updates/sec

Title: US-10-672-787-35_COPY_11357_12736
Perfect score: 1380
Sequence: 1 taaagcgtgaaccatcgttaa .........acgacggcatattctttga 1380
Scoring table: IDENTITY NUC
Searched: Gapop 10.0 , Gapext 1.0
5244920 seqs, 348612423l residues
Total number of hits satisfying chosen parameters: 7439360

Post-processing: Minimum Match 0%
Maximum Match 100%
Minimum DB seq length: 22
Maximum DB seq length: 2000000000
Listing first 45 summaries

Database :    N Geneseq_8:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2004as:*
12: geneseqn2004bs:*
13: geneseqn2005s:*
14: geneseqn2006s:*
15: geneseqn2006s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 1380 | 100.0 | 1398 | 8 | ACA39325 | Aca39325 Prokaryot |
| 2 | 1380 | 100.0 | 96109 | 4 | AAF28548 | Aaf28548 Genomic f |
| 3 | 1370.4 | 99.3 | 1410 | 12 | ADL02987 | Adl02987 DNA encod |
| 4 | 231 | 16.7 | 1344 | 8 | ACA21070 | Aca21070 Prokaryot |
| 5 | 226.2 | 16.4 | 1416 | 4 | ADA29995 | Ada29995 DNA encod |
| 6 | 160.4 | 11.6 | 1314 | 4 | AAS53453 | Aas53453 Haemophil |
| 7 | 160.4 | 11.6 | 1314 | 8 | ACA34366 | Aca34366 Prokaryot |
| 8 | 160.4 | 11.6 | 1314 | 11 | ADL46439 | Adl46439 aspartate |
| 9 | 160.4 | 11.6 | 1314 | 14 | AEC10868 | Aec10868 Haemophil |
| 10 | 160.4 | 11.6 | 2925 | 15 | AEB31251 | Aeb31251 Haemophil |
| 11 | 160.4 | 11.6 | 110000 | 2 | AAT42063_11 | Aat42063_11 Continuation (12 o |
| 12 | 160.4 | 11.6 | 110000 | 2 | AAT42063_12 | Aat42063_12 Continuation (13 o |
| 13 | 157.6 | 11.4 | 110000 | 14 | AEB39175_30 | Aeb39175_30 Continuation (31 o |
| 14 | 157.6 | 11.4 | 14324 | 13 | ADT05533 | Adt05533 Haemophil |
| 15 | 154.4 | 11.2 | 349980 | 13 | ADT05649 | Adt05649 Haemophil |
| 16 | 154.4 | 11.2 | 1281 | 8 | ACA37295 | Aca37295 Prokaryot |
| 17 | 152.8 | 11.1 | 57367 | 14 | AEB39164 | Aeb39164 L. pneumo |
| 18 | 152.8 | 11.1 | 1281 | 8 | | |
| 19 | 152.8 | 11.1 | 110000 | 14 | AEB42737_15 | Continuation (16 o |
| 20 | 152.4 | 11.0 | 1314 | 11 | ADL46441 | Adl46441 aspartate |
| 21 | 152.4 | 11.0 | 1314 | 14 | AEC10870 | Aec10870 Haemophil |
| 22 | 145.2 | 10.5 | 229301 | 14 | AEB35719 | Aeb35719 L. pneumo |
| 23 | 142.2 | 10.3 | 1305 | 8 | ACA42843 | Aca42843 Prokaryot |
| 24 | 139.6 | 10.1 | 1317 | 4 | AAS52295 | Aas52295 E. coli D |
| 25 | 139.6 | 10.1 | 1317 | 8 | AAH81460 | Aah81460 Escherich |
| 26 | 139.6 | 10.1 | 1317 | 8 | ACA32359 | Aca32359 Prokaryot |
| 27 | 132.8 | 9.6 | 2850 | 15 | AEE31236 | Aee31236 Haemophil |
| 28 | 132.8 | 9.6 | 5640 | 15 | AEE31256 | Aee31256 Haemophil |
| 29 | 125.2 | 9.1 | 1323 | 10 | ADF02545 | Adf02545 Bacterial |
| 30 | 123.6 | 9.0 | 1308 | 8 | ACA44759 | Aca44759 Prokaryot |
| 31 | 117.8 | 8.5 | 1182 | 8 | ACA45658 | Aca45658 Prokaryot |
| 32 | 117 | 8.5 | 1347 | 4 | AAS54275 | Aas54275 Pseudomon |
| 33 | 117 | 8.5 | 1347 | 8 | ACA42598 | Aca42598 Prokaryot |
| 34 | 117 | 8.5 | 1371 | 11 | ABD09098 | Abd09098 Pseudomon |
| 35 | 117 | 8.5 | 1401 | 11 | ABD09257 | Abd09257 Pseudomon |
| 36 | 113 | 8.2 | 1450 | 3 | AAZ47131 | Aaz47131 Pseudomon |
| 37 | 113 | 7.9 | 1404 | 8 | ACA53404 | Aca53404 Prokaryot |
| 38 | 111 | 8.0 | 1317 | 8 | ACA51133 | Aca51133 Prokaryot |
| 39 | 110.8 | 8.0 | 4942 | 5 | AAS89816 | Aas89816 DNA encod |
| 40 | 110.8 | 8.0 | 4944 | 5 | AAS89273 | Aas89273 DNA encod |
| 41 | 110.8 | 8.0 | 4944 | 5 | AAS90181 | Aas90181 DNA encod |
| 42 | 110 | 8.0 | 1350 | 8 | ACA43846 | Aca43846 Prokaryot |
| 43 | 109.6 | 7.9 | 1311 | 10 | ACF71682 | Acf71682 Photorhab |
| 44 | 109.6 | 7.9 | 110000 | 10 | ACF67367_48 | Continuation (49 o |
| 45 | 109.6 | 7.9 | 110000 | 10 | ACF65387_0 | Acf65387_0 Photorhab |

ALIGNMENTS

RESULT 1
ACA39325/c/s
ID  ACA39325 standard; DNA; 1398 BP.
XX
AC  ACA39325;
XX
DT  19-JUN-2003  (first entry)
XX
DE  Prokaryotic essential gene #20982.
XX
KW  Antisense; ds; prokaryotic essential gene; cell proliferation;
KW  drug design; gene.
XX
OS  Moraxella catarrhalis.
XX
XX
PN  WO200277183-A2.
XX
PD  03-OCT-2002.
XX
PF  21-MAR-2002; 2002WO-US009107.
XX
PR  21-MAR-2001; 2001US-00815242.
PR  06-SEP-2001; 2001US-00948993.
PR  25-OCT-2001; 2001US-0342923P.
PR  08-FEB-2002; 2002US-00072851.
PR  06-MAR-2002; 2002US-0362699P.
XX
PA  (ELIT-) ELITRA PHARM INC.
XX
PI  Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
PI  Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR  WPI; 2003-029926/02.
DR  N-PSDB; ABU35455.
XX
PT  New antisense nucleic acids, useful for identifying proteins or screening
PT  for homologous nucleic acids required for cellular proliferation to
PT  isolate candidate molecules for rational drug discovery programs.
XX
PS  Claim 14; SEQ ID NO 27195; 1766pp; English.
XX

```
CC  The invention relates to an isolated nucleic acid comprising any one of
CC  the 6213 antisense sequences given in the specification where expression
CC  of the nucleic acid inhibits proliferation of a cell. Also included are:
CC  (1) a vector comprising a promoter operably linked to the nucleic acid
CC  encoding a polypeptide; (2) a host cell containing the vector; (3) an isolated
CC  nucleic acid; (4) an antibody capable of specifically binding
CC  the polypeptide or its fragment whose expression is inhibited by the
CC  antisense nucleic acid; (5) producing the polypeptide; (6) inhibiting cellular
CC  proliferation or the activity of a gene in an operon required for
CC  proliferation; (7) identifying a compound that influences the activity of
CC  the gene product or that has an activity against a biological pathway
CC  required for proliferation, or that inhibits cellular proliferation; (8)
CC  identifying a gene required for cellular proliferation or the biological
CC  pathway in which a proliferation-required gene or its gene product lies
CC  or a gene on which the test compound that inhibits proliferation of an
CC  organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC  compound's activity; (11) a culture comprising strains in which the gene
CC  product is overexpressed or underexpressed; (12) determining the extent
CC  to which each of the strains is present in a culture or collection of
CC  strains; or (13) identifying the target of a compound that inhibits the
CC  proliferation of an organism. The antisense nucleic acids are useful for
CC  identifying proteins or screening for homologous nucleic acids required
CC  for cellular proliferation to isolate candidate molecules for rational
CC  drug discovery programs, or for screening homologous nucleic acids
CC  required for proliferation in cells other than S. aureus, S. typhimurium,
CC  K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC  prokaryotic essential genes. Note: The sequence data for this patent did
CC  not form part of the printed specification, but was obtained in
CC  electronic format directly from WIPO at
CC  ftp.wipo.int/pub/published_pct_sequences
XX
SQ  Sequence 1398 BP; 408 A; 306 C; 295 G; 389 T; 0 U; 0 Other;

Query Match            100.0%;  Score 1380;  DB 8;  Length 1398;
Best Local Similarity  100.0%;  Pred. No. 0;  Mismatches 0;
Matches 1380;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATATAATTTTTAAATTGATC   60
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1383 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATATAATTTTTAAATTGATC 1324

Qy      61 CATGCTTAGCACAAGCAGAGGTGACAAAAAGCACAGCCCTAGCGTGCTTGACTGGCTGCCAA  120
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1323 CATGCTTAGCACAAGCAGAGGTGACAAAAAGCACAGCCCTAGCGTGCTTGACTGGCTGCCAA 1264

Qy     121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT  180
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1263 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 1204

Qy     181 GAGTAAATCTTGCTCAATTATCGGGGCCATCTTGACCGGTAAATAACACTATCCACCCA  240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1203 GAGTAAATCTTGCTCAATTATCGGGGCCATCTTGACCGGTAAATAACACTATCCACCCA 1144

Qy     241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAT  300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1143 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 1084

Qy     301 CAAATGCCAGTGATGATATGCCAAGCCCATCAAATTGCTTGCAATGGTTGA  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1083 CAAATGCCAGTGATGATATGCCAAGCCCATCAAATTGCTTGCAATGGTTGA 1024

Qy     361 GCCGATGGTTGTACCTTTTGAATCATTAAAAATATGCCTTGCTATCATCAATCAATATA  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1023 GCCGATGGTTGTACCTTTTGAATCATTAAAAATATGCCTTGCTATCATCAATCAATATA 964

Qy     421 TTCACAGCCGATGCGCGTAGCCCTTTGAAAGTTCACACACCATGCCATTCCATGGG  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     963 TTCACAGCCGATGCGCGTAGCCCTTTGAAAGTTCACACACCATGCCATTCCATGGG 904

Qy     481 AAGAGTTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCAATTGAGTTAGGTTATGCTTGCC  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 AAGAGTTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCAATTGAGTTAGGTTATGCTTGCC 844

Qy     541 TTTGATTTTTAAATTGATGGATTGGTTAGAAAGTTTTGTTTTTTTTGATGATAATAAATGCCATACC  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 TTTGATTTTTAAATTGATGGATTGGTTAGAAAGTTTTGTTTTTTTTGATGATAATAAATGCCATACC 784

Qy     601 ATCACCCATCATCAAAATAAATAAAAATCGCCGTTTGGGCGGCTAAGTAAGCTATTGTGT  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 ATCACCCATCATCAAAATAAATAAAAATCGCCGTTTGGGCGGCTAAGTAAGCTATTGTGT 724

Qy     661 GGCATCGTCTTGGTAGGATGAGGGCCAACCCCTTGTCAGTGTGTCAAGACAGCATTGCTTGGCTAAATT  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 GGCATCGTCTTGGTAGGATGAGGGCCAACCCCTTGTCAGTGTGTCAAGACAGCATTGCTTGGCTAAATT 664

Qy     721 GGCATCGTCTTGGTAGGATGAGGGCCGGTTTGGCAATTTTCAAAAAAATACGCAAATTTTTGTGC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 GGCATCGTCTTGGTAGGATGAGGGCCGGTTTGGCAATTTTCAAAAAAATACGCAAATTTTTGTGC 604

Qy     781 CAAATAATCCACCATGCCATCGTGCCGATCAAGATAACGGGGAGAGATTTAAAAATGT  840
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     603 CAAATAATCCACCATGCCATCGTGCCGATCAAGATAACGGGGAGAGATTTAAAAATGT 544

Qy     841 TGCCCCTTGAGCTGCCAAAATTTGAAAGATATCTCAAGCTGAAAAACTTGACAAGCAC  900
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 TGCCCCTTGAGCTGCCAAAATTTGAAAGATATCTCAAGCTGAAAAACTTGACAAGCAC 484

Qy     901 CACCAAATCCATATTGGATCGTTAATTAATTCAAGCTGCAGGGCGTGCCAATATTACCGCC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 CACCAAATCCATATTGGATCGTTAATTAATTCAAGCTGCAGGGCGTGCCAATATTACCGCC 424

Qy     961 GACACCAACAATCATGCCTGACAGTGTTTGCCATCTCGCCTACTAAATGTCGTAACAGTGCT 1020
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GACACCAACAATCATGCCTGACAGTGTTTGCCATCTCGCCTACTAAATGTCGTAACAGTGCT 364

Qy    1021 TTTGCGTTTGAGCCTGTGAGGGGCGATGATGGTGTGTTGTTTTGTTCTTGAGTTAGAATCTCG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 TTTGGCATTTGAGCCTGTGAGGGGCGATGATGGTGTGTTGTTTTGTTCTTGAGTTAGAATCTCG 304

Qy    1081 TGCTTTGAGTGTATCAAATGAACAATTGAACATCGCTGATCAACAGAGAAATGCCTTGAGCTTT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 TGCTTTGAGTGTATCAAATGAACAATTGAACATCGCTGATCAACAGAGAAATGCCTTGAGCTTT 244

Qy    1141 GGCGGCGACAATACTCGGGCGTTCTTGGGTTAAATGCCGAGGCTGATGGTGATGATCTGATCGGC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 GGCGGCGACAATACTCGGGCGTTCTTGGGTTAAATGCCGAGGCTGATGGTGATGATCTGATCGGC 184

Qy    1201 TGAGCTTAAGAGTTCACTGTGTCCAACTGCCAAAAAATGACAATGAAAGACGCCCTGTGGGCAGTTT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 TGAGCTTAAGAGTTCACTGTGTCCAACTGCCAAAAAATGACAATGAAAGACGCCCTGTGGGCAGTTT 124

Qy    1261 ATCTGCCAAGGTAGGATTGGGATTACCATGATGATGCTAAACTTTTGTGACCGGTGATTGAC 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 ATCTGCCAAGGTAGGATTGGGATTACCATGATGATGCTAAACTTTTGTGACCGGTGATTGAC 64

Qy    1321 CAAAAAATTCACAGCTGACAGTCAGTCCAGAGCTGCCCAAAACCGACGACGGGCCATATTTTTTTGA 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 CAAAAAATTCACAGCTGACAGTCAGTCCAGAGCTGCCCAAAACCGACGACGGGCCATATTTTTTTGA 4
```

RESULT 2
AAF28548
ID   AAF28548 standard; DNA; 96109 BP.
XX
AC   AAF28548;
XX
DT   04-APR-2001 (first entry)
XX
DE   Genomic fragment #35.
XX
KW   Genomic library; bacteria; human upper airway; otitis media; sinusitis;
KW   bronchopulmonary; endocarditis; meningitis; 88.
XX
OS   Moraxella catarrhalis.
XX

```
PN   WO200078968-A2.
XX
PD   28-DEC-2000.
XX
PF   16-JUN-2000; 2000WO-US016649.
XX
PR   18-JUN-1999; 99US-0140121P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Lagace RE, Patterson C, Berg KL;
XX
DR   WPI; 2001-041427/05.
XX
PT   Genomic library for identifying diagnostic and therapeutic compositions,
PT   and for identifying virulence factors, regulatory elements and drug
PT   targets, comprises Moraxella catarrhalis nucleic acids.
XX
PS   Claim 1; Page 345-368; 545pp; English.
XX
CC   The present invention relates to a Moraxella catarrhalis genomic library
CC   comprising of a combination of 41 nucleic acid molecules (see AAF28514-
CC   AAF28554). The library has a number of uses described in the
CC   specification e.g. is useful for aidentifying diagnostic and therapeutic
CC   compositions. M. catarrhalis (Branhamella catarrhalis) is a large
CC   aerobic, gram-negative diplococcus, normally found among the bacterial
CC   flora of human upper airways. M. catarrhalis is known to cause acute,
CC   localised infections such as otitis media, sinusitis and bronchopulmonary
CC   infection and life-threatening, systemic diseases including endocarditis
CC   and meningitis
XX
SQ   Sequence 96109 BP; 28783 A; 18910 C; 20341 G; 28075 T; 0 U; 0 Other;
```

```
Query Match             100.0%;  Score 1380;  DB 4;  Length 96109;
Best Local Similarity   100.0%;
Matches 1380; Conservative 0;  Pred. No. 0;  Mismatches 0;  Indels 0;  Gaps 0;
```

PD   06-JAN-2004.
XX
PF   04-APR-2000; 2000US-00540236.
XX
PR   08-APR-1999; 99US-0128416P.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Breton GL;
XX
PS   Disclosure; SEQ ID NO 673; 429pp; English.
XX
PT   New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT   preparing a composition for diagnosing, preventing or treating infection
PT   caused by Moraxella catarrhalis.
XX
DR   WPI; 2004-178127/17.
DR   P-PSDB; ADL04907.
XX
CC   The invention relates to an isolated nucleic acid encoding an Moraxella
CC   catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC   composition for diagnosing, preventing or treating infection caused by
CC   Moraxella catarrhalis. The present sequence represents DNA encoding a M.
CC   catarrhalis protein.
XX
SQ   Sequence 1410 BP; 409 A; 307 C; 302 G; 392 T; 0 U; 0 Other;

Query Match   99.3%; Score 1370.4; DB 12; Length 1410;
Best Local Similarity   99.6%;
Matches 1374; Conservative   0; Mismatches   6; Indels   0; Gaps   0;

Qy   1 TAAAGCGTGAAACCATGTTAACAAACTTATCACCCTCCGCTCATTATTAATTTTTAAATTGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1392 TAAAGCGTGAAACCATGTTAACAAACTTATCACCCTCCGCTCATTATTAATTTTTAAATTGATC 1333

Qy   61 CATGTCTAGCACAAGCAGCAGGGCGACAAAAAGCACAGCACGCCCTAGCGGTGCTTGACTGGCA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1332 CATGTCTAGCACAAGCAGCAGGGCGACAAAAAGCACAGCACGCCCTAGCGGTGCTTGACTGGCA 1273

Qy   121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAATGTGCATCACCTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1272 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAATGTGCATCACCTTT 1213

Qy   181 GAGTAAATCTTGCTCAATTATCGGGCCATCTTCTTGACCGATTAAAATACACACTATCCACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1212 GAGTAAATCTTGCTCAATTATCGGGCCATCTTCTTGACCGATTAAAATACACACTATCCACCCA 1153

Qy   241 CTGGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCCAAAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1152 CTGGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCCAAAAT 1093

Qy   301 CAATGCCAGTGATGATAATGCCAAGCCATCAATTGCTGCAATGGTTGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1092 CAATGCCAGTGATGATAATGCCAAGCCATCAATTGCTGCAATGGTTGA 1033

Qy   361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATCATCAATATA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1032 GCCGATGTTTGTACCTTTTGAATCATTAAAATATCATCAATATA 973

Qy   421 TTCACACGGATGGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCACATGGCATCCATGGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   972 TTCACACGGATGGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCACATGGCATCCATGGG 913

Qy   481 AAGATTGGGCAGCTGTGTGCCTAAGGCAGTAAGGCAATTGAGTTATGCTTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   912 AAGATTGGGCAGCTGTGTGCCTAAGGCAGTAAGGCAATTGAGTTATGCTTGCC 853

Qy   541 TTTGATTTTTAATTGAATGGATTGGTAAAAAGCTTGTTTTTTTGATGATATAAATGCCATACC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   852 TTTGATTTTTAATTGAATGGATTGGTAAAAAGCTTGTTTTTTTGATGATATAAATGCCATACC 793

Qy   601 ATCACCCATCAAATAAAATAAAAATCTGCCGTTTGGGTGGGGTCTAAGTAAGCTATTTGTTGTCT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   660

Qy   721 GGCATCGTCGTCTTGGGTAGAATGAATGGGCGGTTTGGGCAATTTCAAAAATACGCAAATTTTTGTGTC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   672 GGCATCGTCTTCTTGGGTAGAATGAATGGGCGGTTTGGGCAATTTCAAAAATACGCAAATTTTTGTGTC 613

Qy   781 CAAATTAATCCACCATTGCCATCGTTCGTCAAGATTGATATCGCCGGCAGGATTTAAAATTGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   732 GATAATAGGTTCGTGATTTTGGGGCAACGGCCTTGTCAGTGTTCAGTGTGGTCAAGATTTGCTTGGCTAAATT 673

Db   612 CAAATTAATCCACCATTGCCATCGTTCGTCAAGATTGATATCGCCGGCAGGATTTAAAATTGT 553

Qy   841 TGCCCCTTTGAGCTGCCAAAATTTGAGATAATGCTCAAGCTGAAAAAACTTGACAGCTCAAGCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   552 TGCCCCTTTGAGCTGCCAAAATTTGAGATAATGCTCAAGCTGAAAAAACTTGACAGCTCAAGCAC 493

Qy   901 CACCAAAATCCATATTTTGATCGTTAATAAATTCAAGTGGGCGGCTGCCAATAATTACCGCGC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   492 CACCAAAATCCATATTTTGGATCGTTAATAAATTCAAGTGGGCGGCTGCCAATAATTACCGCGC 433

Qy   961 GACACCAACAATCAATGCCTGCAGTGTTTGCCATCTCGCCTACTAATGTGTCGTAACAGTGCT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   432 GACACCAACAATCAATGCCTGCAGTGTTTGCCATCTCGCCTACTAATGTGTCGTAACAGTGCT 373

Qy   1021 TTTGGCATTTGAGCCTTGATCGGGTGATGAATGGTTGTTGTTGTTGTTGTTGGTTAGATCTCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   372 TTTGGCATTTGAGCCTTGATCGGGTGATGAATGGTTGTTGTTGTTGTTGTTGGTTAGATCTCG 313

Qy   1081 TGCTTTGAGTGTATCAAATGAACAATTGAACATCGCTTCACAGGATCTGATGGCCTTT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   312 TGCTTTGAGTGTATCAAATGAACAATTGAACATCGCTTCACAGGATCTGATGGCCTTT 253

Qy   1141 GGGCGGCGGACAATACTCGGGCGTTCTTGGGTTAATGCGCGGGGCTAATGATCTGATGGGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   252 GGGCGGCGGACAATACTCGGGCGTTCTTGGGTTAATGCGCGGGGCTAATGATCTGATGGGC 193

Qy   1201 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTCGTGGGGCGTTT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   192 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTCGTGGGGCGTTT 133

Qy   1261 ATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAAACTTTGTGACCGATGATTGAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   132 ATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAAACTTTGTGACCGATGATTGAC 73

Qy   1321 CAAAAAAATTCACAGCTGACAGTCGCCAAACCGACGGGCATAATTTTTTTTTGA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   72 CAAAAAAATTCACAGCTGACAGTCGCCAAACCGACGGGCATAATTTTTTTTTGA 13

RESULT 4
ACA21070/c
ID   ACA21070 standard; DNA; 1344 BP.
XX
AC   ACA21070;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Prokaryotic essential gene #2727.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Acinetobacter baumannii.
XX
PN   WO20027183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PR   21-MAR-2001; 2001US-00815242.

```
PR   06-SEP-2001; 2001US-00948993.
PR   25-OCT-2001; 2001US-03492923P.
PR   08-FEB-2002; 2002US-00072851.
PR   06-MAR-2002; 2002US-03626999P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, ZysKind JW;
PI   Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
DR   WPI; 2003-0229926/02.
DR   P-PSDB; ABU17200.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 8940; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 6213 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation or the activity of a gene in an operon required for
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for cellular proliferation or the biological
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note: The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1344 BP; 385 A; 254 C; 318 G; 387 T; 0 U; 0 Other;

     Query Match        16.7%;  Score 231;  DB 8;  Length 1344;
     Best Local Similarity   52.4%;  Pred. No: 1e-55;
     Matches  721;  Conservative  0;  Mismatches 590;  Indels  66;  Gaps   7;
```

```
Qy    2  AAAGGCTGAAACCATGTTAACAAACTTATCACCTCGGCTCATTATTAATTTTTTAAATTGATCC   61
            |||| |||||     |  ||||| ||| || || |  | ||||||| |  ||||  ||| |||
Db 1340  AACCAAATTCACGCAGGCAACAAAACTGCTGACCACGGTCATTATAAACTTTTAAACATATCA 1281

Qy   62  ATGCTAGACAACAAGCAGGTGACAAAAGCACAGCCCTAGCGCGTCGTCTTTGACTGGCTGCCAAT  121
            ||| |||| |  |  ||||| ||||| ||||  ||      ||||||| | ||| | ||| ||
Db 1280  AAACTTGCACATGCTGGTCGAT------AGCAATAACCACATCTTCAGCTTGTTGTTTCAC 1229

Qy  122  ACTCGCCGATCAAGGTGTACCAAACCTGATATATTTTTTAGATATCCCAAATGTGCATCACCTTTG  181
            || |||| | ||| || ||| |||| | |||| |||| |||| |||| |||  | ||| | ||
Db 1228  GTTGACACAGCTCGACAGCTTCTTTAGGCTGTTGCTCGAGTAAAATACCACTACCCAC 1170

Qy  182  AGTAAATCTTGCTCAAATTATCGGGGGCCATGTCAAATTATTGACCGATTAAAATACCACTACCCAC  241
            |||||||||||||| |  || ||  ||||| || |  ||||  ||| |  |||||| || |||
Db 1169  TGAATGGCTTGCTCGATGACGCGGCCGCCCATCTTCACCAAATCAATACCACAACTTTGGCATAT 1110
```

```
Qy  242  TGCATAACATATGTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCACCCAAAAATC  301
            || ||| | | ||||| |   |  | |||||||||||| | ||||| || || |||||
Db 1109  TTTTCGATAGAGAACGTAAAGGAGAAAAATCTTGACCCCTTTGCCCGTCCAACCTAAAATA 1050

Qy  302  AATGCCCA------GTGGATGAATATGCCCATAAAAACCGGTTGCCAAGCCCATCAATTGCTCCAAGT  355
            | |||  |      |||| || || | |  ||| ||| ||| || ||||| ||| |||| ||
Db 1049  AGTGCGAAACCTTACCTTTTTCACTTCAATGCGCAGCACCTTAGCCATCAATTGCTGTGC 990

Qy  356  GTTGAGCCGATGTTGTACCTTTTGAATCATTAAAAAAATAATGCCTTGCTATCATATCATCA  415
            || || || ||||| ||||||| |||||||| | |  | ||| | ||| | || || | |
Db  989  GTAGCACCAACATTGGTACCTTTAGAGTCATTATAATAACGGACATCACCGCACGTTTTA 930

Qy  416  ATATATTCAGCGATGCGGTAGCCCTGTCCCCCAATGCTAAAACACGCTAAAAGCATTGGCGTTGTGC  475
            ||||| | || |||||  ||| |    ||| | |||||||||||| ||||| ||||| | |||
Db  929  ACGTACTCACAGCCGATGCTCAAATCCTTTAAATCGGTTTTAAGTTTCAAGGCATGACTCC 870

Qy  476  ATGGGAAGAGAATTGGGCAGCTGTTGCCCTAAGGCAAGTGCAGATAAGGCATTGAGTTATGC  535
            ||  |||| | |||| | |||||  |||||||| |||||||||| |||||||| || |||||
Db  869  ATAGGTAAACCAATTGCCTCCCCAAATGCTAAAACACGCTAAAGCCATTGGCGTTGTGC 810

Qy  536  TTGCCTTTGATTTTTAAATTGATGGATTGGGTAAAAAGTTTGTTTTTTTGATGATAATGCC  595
            || || || | ||||||| | ||| |||| |||||| |||||    | ||||| || |||
Db  809  ATACCTTGAATATAATAAAATCCGAGCTTTTAATTAACCGGTTAAACCACACGAGCAAGCCAA 750

Qy  596  ATACCATCACCATCATCAAAATAATTAATAAAATCTGCGTTTTGGCTAGTAGTAAGTCTATTT  655
            ||||||||||||||| |||| |||||||||||||| |||||||||| |||||||| |
Db  749  AGCGGTACCGTCGGCATCTCTTAAAACGCCATATTGATT------TAAATCTGGTGGCGT 698

Qy  656  GTTGTGATTAATGGTGTTGGGGGCCAACCTTGTCAGTGTGGTCAAGCAATTGCTGGGCT  715
            |||   ||||  | ||| || | || | |  | ||| || |    |||||| |||| |
Db  697  TTAAACCAAAGCTTTGCATTGGGTGTCGTACTCGAACAACAGTGGGGCGGCTTA------ 647

Qy  716  AAATTGGCATCGTCTTGTTGTAGAGATGATGGCGGTTTGGCAATTTCAAAAATACGCAATTTT  775
            |      || |||||||| ||  || | || ||||| | |||||| ||||||||||  | |||
Db  646  ------AAGCAATCGTCTCGGGTTAAAATCAACATTTTTTTAGCCGGTGGCTTGGCTATTTTTT 591

Qy  776  TGTGCCAAAATAATCACCACCATCGTGCCGATCAAGATGATCAGCGGAGAGAGATTTAAA  835
            |  |  |||||| || || ||| |||||||| |||| |||||| ||| ||| | ||||||
Db  590  GCTTGGTGATACCCCCAGCATATTTCCATGACGGTCTAAATGGTCTTCACTCATATTGAGA 531

Qy  836  ATTGTTGTGCCCCTTTGAGCTGCCAAATTTGAGATATGTCAAGCTGAAAAACTTGACAGCTCA  895
            | | || |  || |||||| || ||||||||||| ||| |||||||| |||| |||||| | |
Db  530  ACCACTGCTACCTCAGCCGTTTAAGTAGGAGGTGGTTTCTTAAGCTTGATAAGTCA 471

Qy  896  AGCACCACCAAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGGTGCCCAATATTA  955
            |||||||   |||| |||| || || |||| ||||||| ||| | |||| | ||||| ||||
Db  470  AGGACTAGTAGCTC---TGGTGATCTTTAAGTAAATCTAACCGCGGTCGACCCAAGATTG 414

Qy  956  CCGCCGACACCAACAAATCATGCCCTTGCATGTTTTGCCACTACTAATGTCGTTAACA  1015
            |||||  |||||| ||| ||||||| |||| |  | ||||| || |||||||||| ||
Db  413  CCGCCTACTCGACAACTTTCTTTACCCTGCATCCTTAGCCATTAAAGTCGTTACA 354

Qy 1016  GTGCTTTTGGCCAATTTGAGCCTGTGATGGGCCGATGAGTTGTTGTTTGGTTGCTTGTTAGA  1075
            |  ||||| ||| |||| ||| ||| || |||| |  |   ||| ||| |||  |||| | |
Db  353  GTACTTTTTGCCATTGGAACCTGTAATCGGCCACAGACACATCAGTGGGC-------- 304

Qy 1076  TCTCGTGCTTTGAGTTGTATCTCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTTTA  1135
            | | |  |  | ||| | || |  || ||| ||| ||| | || | ||| ||||||| ||||||
Db  303  -------------ACGACGCAATAATTGGATAATCACCCACCACAGAAATACCTTTA 261

Qy 1136  GCTTTGGGCGCGACAAATACTCGGCGTTCTTGGGTTTAATGCGCGGGCGCTAGTCTGA  1195
            |  | ||| || ||||||||| ||  ||||| || | || |||  | || | ||| |
Db  260  GCAATAGCTGCCTGATTTTCCGGTAATTGTCGTGCAAGCCGGCTTAAATAATCTCT 201

Qy 1196  TCGGCGCTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAAGCGCTCGGGGC  1255
            ||  | |||| ||||| ||||| ||| ||||||| | |||||| || || |  | ||| | |
Db  200  TCTGCTTGTTAATAAAATTCTTGATCAAGCTGACCAAAAACTGGGTTTTAAGGCCACCGGA 141

Qy 1256  AGTTTATCTGCCAAGGTAGGATTGGGATTACCATCCGATGATGCTAACTTTGTGACCGTGA  1315
            | ||| ||| | || || | || | | || ||  ||| ||||| ||  || || | |||||
Db  140  ATCTGATCGTGTTCCGGGAGGTGTGGGGGGTCCGTTACAGCAACTTGGTAGCCTTGT 81

Qy 1316  TTGACCAAAAAAATTCACAGAGTTGACAGTCCAGAGCGTGCCAAAACCGACGACGGGCATAT  1372
            || |||||||||||| || | | | ||||| || |||||| || |||| || | || ||||
Db   80  ... GGCATAT 1372
```

RESULT 5

AC ADA29995;
XX
ID ADA29995/c standard; DNA; 1416 BP.
XX
AC ADA29995;
XX
DT 20-NOV-2003 (first entry)
XX
DE DNA encoding Acinetobacter baumannii protein #1282.
XX
KW ds; gene; Acinetobacter baumannii; bacterial disease; antibacterial;
KW vaccine; plant biocontrol agent.
XX
OS Acinetobacter baumannii.
XX
PN US6562958-B1.
XX
PD 13-MAY-2003.
XX
PF 04-JUN-1999; 99US-00328352.
XX
PR 09-JUN-1998; 98US-0088701P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Breton G, Bush D;
XX
DR WPI; 2003-576092/54.
DR P-PSDB; ADA34121.
XX
PT New Acinetobacter baumannii proteins and nucleic acids, useful as reagents
PT for diagnosing a bacterial disease, as components of antibacterial
PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for
PT plants.
XX
PS Example; SEQ ID NO 1282; 328pp; English.
XX
CC The invention relates to isolated Acinetobacter baumannii nucleic acids.
CC The A. baumannii nucleic acids and polypeptides are useful as reagents
CC for diagnosing a bacterial disease, as components of antibacterial
CC vaccines, as targets for antibacterial drugs, to detect the presence of
CC A. baumannii and other Acinetobacter species in a sample, in screening
CC compounds for the ability to interfere with the A. baumannii life cycle
CC or to inhibit A. baumannii infection, and as biocontrol agents for
CC plants. The present sequence represents DNA encoding an A. baumannii
CC protein.
XX
SQ Sequence 1416 BP; 417 A; 262 C; 326 G; 411 T; 0 U; 0 Other;

Query Match   16.4%;   Score 226.2; DB 9; Length 1416;
Best Local Similarity   52.1%;   Pred. No. 2.5e-54;
Matches 718;   Conservative   0; Mismatches 593;   Indels 66;   Gaps   7;

Qy   2 AAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTCATTAAATTGATCC 61
              ||| | || | |||| | || || | ||||||| | |||| | |||| ||||
Db   1409 ACGAATTGACGCAGGTGACAAAAGCACAGCCCTAGCGCGTCATTATAACTTTTAAACATATCA 1350

Qy   62 ATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGCGTGCTTGCTTGACTGGTCGATCTGA 121
              |||||| || | |||||| || ||||| || || ||||| ||| |||||
Db   1349 AAACTTGCACATGCTGGTGAT------AGCAATACCACACATCTTCAGTTGTGTTTCAC 1298

Qy   122 ACTGCCCGATCAAGTGTACCAACCTGATATATTTAGATGCCAAAATGTGCATCACCTTTG 181
              ||||| ||||| | | ||| || |||| |||||| || |||| | ||||||||||
Db   1297 GTTGACACAGCTCGACAGCTTCTTTAAGCGTTGCTGCATGTAAATATACACTTATCCAC 1239

Qy   182 AGTAAAATCTTGCTCAAATTATCGGGGCCATCTTCACCAATCAATACCACAACTTTGGCATAT 241
              |||| | ||| | |||| | | ||||| | || ||||||||| | | || ||||
Db   1238 TGAATGGCTTGTTCGATGACAGGGCGCCATCTTCACAGAGCGTGCCCAAAACCGACGACGGGCATAT 1179

Qy   242 TGCATAACATATGCTGACAGCTCACTACTAAAAATCTGACCTTTACCAAGCCCAAAATC 301
              | || | ||||||||| || || | | ||| ||||| ||||| | ||| ||||
Db   1178 TTTTCGGATAGAGAACGTAAAGGAGGAGAAAAAATCTTGTCCTTTGCCCTGTCCAAGT 1119

Qy   302 AATTGCCA------GTGATGAATATGCCATAAAACCGTTGCCAAGCCATCAATTGCTGCAATG 355
              |||||||| || | | ||| || ||||| | || | | |||| | | | |
Db   1118 AGTTGCAAACTTACCTTTTTCACTTTTTAAACATGCTTATCAAGTCCAAGT 1059

Qy   356 GTTGAGCCGATGTTTTGTTACCTTTGAATCATTCATTAAATCTTGCTATCAATATCATCA 415
              | || || || |||| ||| ||||||| ||||| || | || || | ||||| |
Db   1058 GTAGCACCAACATTAGTAGTACCTTTAGAGTCATTAATAACGGACATCATGCACGGTTTA 999

Qy   416 ATATATTCACAGCGATGCGGTAGCCCGATGGGCGGTAGTAGCTTTGAAAGTGCATCC 475
              | | || || || | || |||| || |||||| || | ||||| | || ||
Db   998 ACGTACTCACAGCCGTCGCCCTAGGCCTTTAGAGTCATTAAAACCACATGCATGGTAGC 939

Qy   476 ATGGGAAGATTGGCAGACTGTGGCCTTAAGGCAAGTGCAGATAAGGCATTGACTGAGTTTATGC 535
              | || | | | || || | || | | ||| | | | | | || ||| | | ||
Db   938 ATAGGTAAACCAATTGCTTCCCCAATGTCAAACGATAAAGAGATTAGCTACATTGTGC 879

Qy   536 TTGCCTTTGATTTTGATTGAATGGATTGGTAAGAAGTTGTTTTTTTTGATGAATAATGCC 595
              || | |||| |||||| | ||| ||| || ||| | |||||||| |||| | |
Db   878 ATACCTTTGAATATAATATAGAGCTTTTAATTAACGGTATCATTGTTTTTTTAGTTTATTT 819

Qy   596 ATACCATCACCATCATCAAATAAAATCTGCCGTTTGGTGGCTAAGTAAGTCTATTT 655
              |||||| | | || || | |||||| || | || || |||| |||| || |
Db   818 AGCGTACCGTCGGCATCTCTTAAAAAGCCCATATATGAATT-----TAAATCCGGTGGCGT 767

Qy   656 GTTGTGATAATGGTGTGATTTGGGGCAAGCCTTGTCAGTGTGGTCAAGCAATGGCT 715
              | | | | |||| ||| |||| || || | | |||||| |||| | |||
Db   766 TTAAACCAAAGCTTTGACATTGGTGTCGTATCGTGGAACAAGTGGACAGGCCTTA------- 716

Qy   716 AAAATTGGCATCGTCTCTTGGTGAGAGATGATGGCAATTTGCAAATAATACGCAATTTT 775
              | |||| || | |||| ||| | | |||||||| || |||||||| | | ||||
Db   715 ----ACGACATCTCGGTTAAATACAACTTTTTAGCGCCTTGAAAATACGGATGTTTT 660

Qy   776 TGTGCCAAATAATCCACCATGCCATCGTGCCGATCGAAGATGAGATCAGCGGAGAGATTTAAA 835
              | | ||| || || ||| ||||||| ||||||| || || | | | || | |||||
Db   659 GCTTGATGATACCCCAGCATATTTCCATCAACGGGTCTTAAATGGTCTTCACTCATATTGAGA 600

Qy   836 ATTGTTTGCCCCTTGAGCTGCCAAATTTGACATATGCTCAAGCTGAAAAACTTGACAGCTCA 895
              | || | | |||||||| || | || |||| | || ||| || | ||| || ||
Db   599 ACCACTGCTACCCTCAGCCGTTAAGTACGAGAGGTGGTTTCTAACTGAAAGCTTGATAACTCA 540

Qy   896 AGCACCACCAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCCGGTGCCAATATTA 955
              || ||||||| ||||||| | ||| ||||| || | || | |||||| || | ||
Db   539 AGGACCAGTAACTCTGG---TTGACTTTAAGTAAATCTAACGGCTGGTGTCGACCAAGGTTG 483

Qy   956 CCGCGGACACCAAAATCATGCCTGCACCGATGTTTTGCCATCTGCCGCTTCAATTAACA 1015
              |||| | ||| | || || | || | || ||||| ||| || ||| | ||||
Db   482 CCGGCTACTGCCAACTTCTTACCTTAGCCCATTCCTTAGCCAATTGAAAGTGGTTTACA 423

Qy   1016 GTGCTTTCTTGGCATTTGAGCCTGTGATGGCCATGATTGGTGTGTTTGTTCTTGAGTTAGA 1075
              | | |||| || | | || | ||| | |||| | |||| | | | ||| | |
Db   422 GTACTTTTTCCATTGGACATGCGTTAATCGCCCACACATCAGTGGGC----------- 373

Qy   1076 TCTCGTGCTTTGAGTGTATCAAATGAACAATGAACATCGCGATCACAGAGAATGCCTTGA 1135
              ||| |||| || | | | | |||| | || | ||| |||| || || |||
Db   372 --------ACGGGCAATAACTGTGATATCGCGACCACGGGCACAGAATAACCTTTA 330

Qy   1136 GCTTTGGCGGCGGGGACAAATACTCGGGGCGTTCTTGGGGTTAAATGCCGGGGCTGATGATCTGA 1195
              | ||||| | || ||| | ||| | || |||| | || || || |||| || | | | |||
Db   329 GCAATAGCTGCCTGAATTGTCCGGTAATTGTTGCAAGGCCTTAAAAATAATCTCT 270

Qy   1196 TCGGCTGAGCTTAGAGTTCACTGTCCAAGCTGCCAAATGACAATGAAGCCGTTGGGC 1255
              | |||||| || | | ||| | ||| ||| | | || | |||| ||| ||||
Db   269 TCTGCTTGTGTAATAATAACTTGGTGTTAATAATGGTTTTAAACAAGGCCTCTGGGC 210

Qy   1256 AGTTTAATCTGCCCAAGGTTAGGGATTAGGGATTACCATCGATGATGAGCTAACTTTGTTGACCGGTGA 1315
              | || | |||| ||| | || | ||| | | || ||||| | ||| ||| || | |||| ||
Db   209 ATCTGATCGTTCCGGACGAGGTGTGGGGGAATCCGTTACGACAACTTGGTAGCCGTTGT 150

Qy   1316 TTGACCAAAAAATTCACAGCTGACAGTCGCCCAAACCGACGACGGGCATAT 1372
              || ||| | | || || || ||| |||| | | |||| || || |
Db   1316 TTGACCAAAAAATTCACAGCTGACAGTCGCCCAAACCGACGACGGGCATAT 1372

```
Db   149 TCATGCAGGAAATTACAGCAGAAACACCTGATTATTCCCAAGCCTGCTACAACTTTT 93
             |  | |||||||    ||  | |||  | | | ||  | |||  || |  || | |
Qy     1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATTAATTTTAAATTGATC 60
             |||| ||||||  || |  || ||| |||||| ||  | || ||| || |||| ||||
```

```
Query Match             11.6%;  Score 160.4;  DB 4;  Length 1314;
Best Local Similarity   49.6%;
  Matches  683;  Conservative    0;  Mismatches 606;  Indels  89;  Gaps  7;
```

```
Db  1308 TAAACATTGAGCTAAAATGGGTAAATTCTTCGCCCGCGTTTCAAAGAGAGCAAACTGATC 1249
              |||  ||    ||  ||||| |||| || | ||| | |||||| ||| |||  |||||
Qy    61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCCTAGCGGTGGCTTTGGGCTGCCAA 120
```

```
Db  1248 GAGACTTGCACAAGGGGGACAAGATAAATAACCA------TATCTCCGCTTTGCAA 1201
              |||||  | |||| | || |||  ||  |||||       || ||| | |||||
Qy   121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTAGATGCATCACCTTT 180
```

```
Db  1200 TGTTGGGCGGTAAAAATTCTATGCGTTGTTCCATTGTATGCAGCAA------ATAACTTG 1147
              |||||| |    ||||  || |   | |  |  || || | |||        ||| ||
Qy   181 GAGTAAAATCTTGCTTCAATTATCGGGGCCATCTTGACGGATTAAAATACACACTATCCACCCA 240
```

```
Db  1146 CGATGAAAAATTTTGCAAGCAGCGCACCACCTCGACCACAACAATAACAAATAAGTGTGG 1087
              || |||||  || ||||||||   || |||| || || |||||| ||  |||| |||
Qy   241 CTGCAATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCACCCAAAAT 300
```

```
Db  1086 TTGATTAATTAATTCAGCTAAATTCTGAAAAATCAGCCCCTTTTCCGTCTCCGGCCTAGCAA 1027
              |||| || ||  |||||||  |  |||| ||||| | ||  ||| ||| || | | |||
Qy   301 CAATGCCACGTGATGATATATCCATAAACCGTGCCAAGCCCATCAATTGCTGCCAAATGGTTGA 360
```

```
Db  1026 CAAATGCAATTACCCTCAATATA--------AAGCCCAGCCAATGCAGCAACTGTACT 976
              || |||||||  |  ||  |||        || ||||   ||| |    ||| |||
Qy   361 GCCGATGTTGTACCTTTTGAATCATTAAAAATATGCCTTGCTATCAATATCATCAATATA 420
```

```
Db   975 CCCCACATTGTTGCTTTAGAGTCATTATTCAACGAATGCCATTAGCTTGATGCACTAA 916
              |||| |||||| |  ||||| ||| || |||| ||| ||| || ||| || |||||
Qy   421 TTCACAGCGATGCGGTAGCCCCTTTGAAAGTCTTAAGGGGTATCAAGCCATGGCATGGG 480
```

```
Db   915 TTGAAAACGATGAATCTAACCCTTTGAAAATGACGGAAGGTGCGGTACGGTAGTCTAAATT 856
              ||||| || |||  | ||  | || |||| | |||||| || ||  | | ||  |||
Qy   481 AAGATTGGGCAGCTGTGCCTAAGGCAAGTCCAGATAAGCATTGAGTAGGTTTATGCTTGCC 540
```

```
Db   855 AATACCTATAGCTTGTGCCAAATGCTTGTGCTGCCAAAAGTTCATTATATAATTATGCGACC 796
              ||| |||| || || ||| || ||  | ||||  |||||| || || |||  | | | ||
Qy   541 CGCTTGATTTTTCGTTTTCCCCAAAAGTCAGCC-------- 600
```

```
Db   795 AACCAAATGTAGCTTCTTCACAAGGTAAAATCACTTCATCTTTTTACCATTAAATATAGTGCTT 736
              |||  || || |  || || |  |||| ||| ||| | |  | | ||| |||||  | |||
Qy   601 ATCACCATCATCAAATAAAATAAAAATCTGC--CGTTTGGTGGCTAAGTAAGCTATTGTTG 659
```

```
Db   735 GCCATTTCAGTTTTTTAGCCAAATACCGGCAATAACTCCGAAAATAGAAAACGGGTAAAATG 676
              |||||| |||||||| |||| || | || |||  ||| | || |||| || || |||
Qy   660 TGATAAATGGTGTGAATTTGGGGCAACAGCCTGTCAGTGTGGTCAACGCATTGCTTGGCTAAAT 719
```

```
Db   675 CGCTTGATTTTCGTTTTCCCCAAAAGTCAGCC-------- 644
              |  | |  ||| ||||| |||| |||||||||
Qy   720 TGGCATCGTCTTGGTAGATGAATGGCGGGTTGGGCAATTTCAAAAATACGGAATTTTGTG 779
```

```
Db   643 ---TATCTTCATTGTTCAACACACTACTTAGCATTATGATAAGGAAAGCTTCGTAAAATTTTGCTT 587
                 | || ||  || ||| |  |  |  || || || ||| || ||| ||| | ||| | |||
Qy   780 CCAAATAATCCACCATGCCGATCCAAGATGACTCAGCCGGAGGAGATTTAAAATTG 839
```

```
Db   586 GGGGATAAATCTTCTAAAATCCATATAGCGGATCCATATAGATCTTCAGTTCAGCACAG 527
              |||||  | |||  || ||| ||| ||| || | || || |||| ||| | || |||
Qy   840 TTGCCCCTTGAGCTGCCAAATTTGAGATAAGCGTCAAGCGTGAAAACTTGACAGCCTCAAGCA 899
```

```
Db   526 TCGCTGCCGCCGAGCTTTTAGGCTATAAGTGTCTCAAGCTAAAAGCTCTAGTA 467
              || |||| | || |||||| || |||||||| | || | ||||| || |||||
Qy   900 CCACCAAATCCATATTTTGGATCGTTAATATAATTCAAGTGGCAGGGGTGCCAATTATTACCGC 959
```

```
Db   466 CATAAAAGTTCACAATCTTCAATTC---AACAATGACAAAGCGGGAAATCCCAAATATTTCCGC 410
              |||| ||| ||| |  |||| |||    ||| |||| | | || |  ||  |||| | ||||
Qy   960 CGACACCAACAAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAAAGTCTAACAGTGC 1019
```

```
Db   409 CCATACCAACTTTCACCAGCAGCTTTCGCCATTTCATTAAGTAGTTTACGGTAC 350
              || || || |||||| |||| ||||| ||||| |||  |  |||| || || ||
Qy  1020 TTTTGGCATTTGACCCTGTGGATGGCGATGTTGTGTGTTGTGCTTGAGTTAGATCTC 1079
```

```
Db   349 TTTTACCCATTTGAACCTGTAATCCCCACAATTGGCTTGTTCGCGCGCGCA----- 298
              |||| ||| |||| | ||| |  | || ||| |  || ||  | | || || |
Qy  1080 GTGCTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGGAATGCCTTGAGCTT 1139
```

```
Db   297 ----------------GAATAATTCAATATGCGCGATTACTTCCACTCCGCTTAA 257
                               || ||| |||| ||| ||| | | || || | ||||
Qy  1140
```

RESULT 7

ACA34366/c

```
ID   ACA34366 standard; DNA; 1314 BP.
XX
AC   ACA34366;
XX
DT   19-JUN-2003 (first entry)
XX
DE   Prokaryotic essential gene #16023.
XX
KW   Antisense; ds; prokaryotic essential gene; cell proliferation;
KW   drug design; gene.
XX
OS   Haemophilus influenzae.
XX
XX
PN   WO200277183-A2.
XX
PD   03-OCT-2002.
XX
PF   21-MAR-2002; 2002WO-US009107.
XX
PI   Wang L,       Zamudio C,     Malone C,      Haselbeck R,   Ohlsen KL,   Zyskind JW;
PI   Wall D,       Trawick JD,    Carr GJ,       Yamamoto R,    Forsyth RA,  Xu HH;
PA   (ELIT-) ELITRA PHARM INC.
XX
PT   New antisense nucleic acids, useful for identifying proteins or screening
PT   for homologous nucleic acids required for cellular proliferation to
PT   isolate candidate molecules for rational drug discovery programs.
XX
PS   Claim 14; SEQ ID NO 22236; 1766pp; English.
XX
CC   The invention relates to an isolated nucleic acid comprising any one of
CC   the 613 antisense sequences given in the specification where expression
CC   of the nucleic acid inhibits proliferation of a cell. Also included are:
CC   (1) a vector comprising a promoter operably linked to the nucleic acid
CC   encoding a polypeptide whose expression is inhibited by the antisense
CC   nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC   polypeptide or its fragment whose expression is inhibited by the
CC   antisense nucleic acid; (4) an antibody capable of specifically binding
CC   the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC   proliferation; (7) identifying a compound that influences the activity of
CC   the gene product or that has an activity against a biological pathway
CC   required for proliferation, or that inhibits cellular proliferation; (8)
CC   identifying a gene required for proliferation or cellular
DR   WPI; 2003-0229926/02.
DR   P-PSDB; ABU30496.
CC   pathway in which a proliferation-required gene or its gene product lies
CC   or a gene on which the test compound that inhibits proliferation of an
CC   organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC   compound's activity; (11) a culture comprising strains in which the gene
CC   product is overexpressed or underexpressed; (12) determining the extent
CC   to which each of the strains is present in a culture or collection of
CC   strains; or (13) identifying the target of a compound that inhibits the
CC   proliferation of an organism. The antisense nucleic acids are useful for
CC   identifying proteins or screening for homologous nucleic acids required
CC   for cellular proliferation to isolate candidate molecules for rational
CC   drug discovery programs, or for screening homologous nucleic acids
CC   required for proliferation in cells other than S. aureus, S. typhimurium,
CC   K. pneumoniae or P. aeruginosa. The present sequence is one of the target
CC   prokaryotic essential genes. Note:The sequence data for this patent did
CC   not form part of the printed specification, but was obtained in
CC   electronic format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1314 BP; 403 A; 214 C; 294 G; 403 T; 0 U; 0 Other;
```

Query Match   11.6%; Score 160.4; DB 8; Length 1314;
Best Local Similarity   49.6%; Pred. No. 2.5e-35;
Matches 683; Conservative   0; Mismatches 606; Indels 89; Gaps 7;

```
Qy    1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATATTTTAAATTGATC 60
          |||  |     || ||| | ||  |    | | | | ||| | |||| |||| |||||
Db 1308 TAAACATTGAGCTAAAATGCGTAAAATCTTCGCCGGCTTTCAAAAGAAGCAAACTGATC 1249

Qy   61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTGACTGGGCTGCCAA 120
          |    |  ||||||||| ||| ||      |||  ||  || |||| ||      |||||
Db 1248 GAGACTTGCACAAGCAGGCGACCAATAATAACCA------TATCTCCGGCTTTGCAA 1201

Qy  121 TACTGCCCGATCAAGTGTACCAAACCTGATATATATTTTAGAGTGCCAAAATGTGCATCACCTTT 180
Db 1200 TGTTGGGCCGTGAAAAATTCTATCGCTTGTTCCATTGTATCGGAACCAA------ATAACTTTG 1147

Qy  181 GAGTAAAATCTTGCTCAAATTATCGGGGCCATCTTGACCGATTAAAATACACTATCCACCCA 240
Db 1146 CGATGAAAAATTTGCAAGCAGCCACATCTCGACCAAAACAATAACAAATATAATGTGTGG 1087

Qy  241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
Db 1086 TTGATTAATTAATTCAGCTAAATTCTGAAAAATCACCCCTTTTCCGTCTCCGGCTTAGCAA 1027

Qy  301 CAAATGCCAGTGATGATGATATGCCTAAGGCAAGCTGCCAGATAAGGACCATCAATTGCTGCAATGGGTTGA 360
Db 1026 CAAATGCAATTTACCCTCAATATA--------AAGCCCAGCCAATGCAGCAACTGCTACT 976

Qy  361 GCCGATGTTTTGTACCTTTTGAATCATTAAAAATGCCCTTGCTATCGAATCATCAATATA 420
Db  975 CCCCACATTGTTGCTTAGAGTCATTAATCCAACGAAATGCCATTAGCTTGATGCTGACTAA 916

Qy  421 TTCACGCCGATGCGCGTAGCCCTTTGAAAAGTCTTAAGGGTATCAAGCATGGCATCATGGG 480
Db  915 TTGAAAACGATGAATCTAACCCTTTGAAAATGACGAAGTGCGGTACGAAATGAAATCTAAAT 856

Qy  481 AAGATTGGGCAGCTGTGCCTAAGGCAAGTGCCAGATAAAGACGAAGGTAGGTTATGCTTGCC 540
Db  855 AATACCTATAGCTTGTGCCAAAGTTCATTATAATTATGCGACC 796

Qy  541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTGATGATAAAATGCCATACC 600
Db  795 AACCAATGTAGCTTCTTCACAAGGTAAAATCACTTACCATTAAATATATGCTT 736

Qy  601 ATCACCCATCATCAAATAAATAAAAATCTCGC--CGTTTGGTGCTAAGTAAGCTATTGTTG 659
Db  735 GCCATTTCAGTTTTAGCCAATAATCCGCACTATTTCCGGAAAAAGAAACGGTATGGTTT 676

Qy  660 TGATAATGGTGTGATTGGGGCAACGCCTGCTCAGTGGTGGTCAAGCATTGCTGGCTAAAT 719
Db  675 CGCTTGATTTCGTTTCTTTCCCAAAAGTCAGCC-------------- 644

Qy  720 TGGCATCGTCTTGGGTAGATGATGGCGGGTTTGGCAAATTTCAAAAAATACGGAATTTTGTG 779
```

```
Qy 1140 TGGCGCGGCGACAATACTCTCGGCGTTCTTTGGGGTTATGCCGGGGGCTGATGATGATCTGATCGG 1199

Qy 1200 CTGAGCTTAAGAGTTCACTGCTCCAAGCTGCCAAAATGACAATGAAACGCCTGTGGGCAGTT 1259

Qy 1260 TATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGGTCTAACTTTTGTGTACCGTGATGA 1319

Qy 1320 CCAAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACCGACGGCATATTTTT 1377

Db   76 AGAGATAATCCACACAAGAAAGACCTGTTTTGCCAAGCCCGATGATAGTAATATTTTT 19
```

```
RESULT 8

ID   ADL46439 standard; DNA; 1314 BP.
XX
AC   ADL46439;
XX
DT   20-MAY-2004  (first entry)
XX
DE   aspartate semialdehyde dehydrogenase DNA #1.
XX
KW   ds; gene; antibacterial;
KW   UDP-N-acetylglucosamine 1-carboxyvinyl transferase-1;
KW   CTP:CMP-3-deoxy-D-manno-octulosonate transferase;
KW   UDP-N-acetylmuramylalanyl-D-glutamate-2,6-diaminopimelate ligase;
KW   D-alanine adding enzyme; D-alanine-D-alanine ligase;
KW   UDP-N-acetylpyruvoylglucosamine reductase;
KW   UDP-N-acetylglucosamine pyrophosphorylase;
KW   UDP-N-acetylmuramoylalanine-D-glutamate ligase;
KW   DP-N-acetylmuramate:alanine ligase; aspartate semialdehyde dehydrogenase;
KW   UDP-N-acetylmuramoylalanyl-D-glutamate: X-ray diffraction analysis.
XX
OS   Haemophilus influenzae.
XX
PN   WO200308735-A2.
XX
PD   23-OCT-2003.
XX
PF   08-APR-2003; 2003WO-CA000481.
XX
PR   08-APR-2002; 2002US-0370899P.
PR   08-APR-2002; 2002US-0370915P.
PR   09-APR-2002; 2002US-0371107P.
PR   09-APR-2002; 2002US-0371185P.
PR   31-MAY-2002; 2002US-0385426P.
PR   06-JUN-2002; 2002US-0388283P.
PR   01-AUG-2002; 2002US-0403348P.
PR   06-NOV-2002; 2002US-0424395P.
PR   08-NOV-2002; 2002US-0425200P.
PR   24-DEC-2002; 2002US-0436345P.
PR   24-DEC-2002; 2002US-0436349P.
PR   26-DEC-2002; 2002US-0436568P.
PR   27-DEC-2002; 2002US-0436675P.
PR   27-DEC-2002; 2002US-0436734P.
PR   27-DEC-2002; 2002US-0436885P.
PR   27-DEC-2002; 2002US-0436885P.
PR   27-DEC-2002; 2002US-0436889P.
PR   27-DEC-2002; 2002US-0436893P.
PR   27-DEC-2002; 2002US-0436900P.
PR   30-DEC-2002; 2002US-0437013P.
XX
PA   (AFFI-) AFFINIUM PHARM INC.
XX
PI   Edwards A, Dharamsi A, Veddadi M, Domagala M, Houston S, Avrey D;
PI   Beattie B, Mansoury K, Ouyang H, Vallee F, Richards D, Nethery K;
PI   Virag C, Buadzija K, Pinder B, Alam MZ, Tai M, Canadien V;
PI   Kanagarajah D, Thalakada R;
XX
DR   WPI; 2003-865361/80.
DR   P-PSDB; ADL46440.
XX
PT   New recombinant bacterial enzymes involved in cell membrane biogenesis,
PT   useful for designing potential antibacterial agents.
XX
PS   Claim 467; SEQ ID NO 157; 407pp; English.
XX
CC   The invention relates to isolated, recombinant polypeptides (I) that have
CC   at least one activity of specified bacterial enzymes involved in cell
CC   membrane biogenesis. (I) are: UDP-N-acetylglucosamine 1-carboxyvinyl
CC   transferase-1 of Streptococcus pneumoniae (S.p), Pseudomonas aeruginosa
CC   (P.a.) or Staphylococcus aureus (S.a.); CTP:CMP-3-deoxy-D-manno-
CC   octulosonate transferase of Escherichia coli (E.c.) or Haemophilus
CC   influenzae (H.i.); UDP-N-acetylmuramylalanyl-D-glutamate-2,6-
CC   diaminopimelate ligase of P.a.; D-alanine-D-alanine adding enzyme of S.a.
CC   or P.a.; D-alanine ligase of Enterococcus faecalis (E.f.); UDP-N-
CC   -acetylpyruvoylglucosamine reductase of P.a. or H.i.; UDP-N-
CC   acetylglucosamine pyrophosphorylase of E.f. or H.i. or S.a.; UDP-N-
CC   acetylmuramoylalanine-D-glutamate ligase of E.f. or H.i.; DP-N-
CC   acetylmuramate:alanine ligase of E.c.; and aspartate semialdehyde
CC   dehydrogenase of H.i and UDP-N-acetylmuramoylalanyl-D-glutamate (sic) of
CC   H.i. Crystalline (II) are used to determine (by X-ray diffraction
CC   analysis) the structural coordinates of (I), and these then used to
CC   design modulators of (II), potential therapeutic agents for treating
CC   diseases caused by the specified bacteria. This sequence represents a DNA
CC   of the invention.
XX
SQ   Sequence 1314 BP; 403 A; 214 C; 294 G; 403 T; 0 U; 0 Other;

Query Match             11.6%;  Score 160.4; DB 11;  Length 1314;
Best Local Similarity   49.6%;
Matches  683; Conservative   0; Mismatches  606;  Indels  89;  Gaps   7;

Qy          1 TAAGCGTGAACCATGTTAACAAACTTATCACCTTGCTCATTATAATTTTAAATTGATC 60
              ||||    ||| | ||  ||| |  ||  ||| || |||  ||||||||   |||||||
Db       1308 TAAACATTGAGCTAAAATGCGTAAATTCTTCGCGCCGGCTTTTCAAAAGAGACAAACTGATC 1249

Qy         61 CATGCTAGCACAAGCAGGTGGAACAAGAGCACAGCCCTAGCGGTGCTTTGACTGGGTGCCAA 120
              |||   |||   ||   || ||||| ||   ||||||   |      |||||||||||||||
Db       1248 GAGACTTGCACAAGCAGGCGGAACAATAATAACCA------------TATCTCCGCTTTGCAA 1201

Qy        121 TACTGCCCGATCAAGTGTACCAACCTGATATATTATTTTTAGATGCCAAAATGTGCATCACCTTT 180
              ||  ||  | || |||| ||| |||  |||||| || ||   |||| ||||||   || ||||||
Db       1200 TGTTGGGCGTAAAAATTCTATCGCTTGTTCCATTGTTATCGGAACAA----------ATAACTTTG 1147
```

```
Qy    181 GAGTAAATCTTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACTATCCACCCA 240
              |||||    |||| || |||  ||| || |||| || | |||| || ||  ||
Db   1146 CGATGAAAAATTTTGCAAGCAGGCGACCAATCTCGACCAAAAATAATAATGTGG 1087

Qy    241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
              | ||||||  ||| ||||||||| ||||||||| |||||||||||||||  || ||
Db   1086 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCCGTCTCCTAGCAA 1027

Qy    301 CAATGCCAGTGATGATATGCCATAAAACGTGCCAAGCCCATCAATTGCTGCCAATGGTTGA 360
              || |||||||||| |||||||| ||| || ||||||||||||| |||| ||||  ||
Db   1026 CAAATGCAATTACCCTCAATATA--------AAGCCCAGCCAATGCAGCAACTGTACT 976

Qy    361 GCCGATGTTTTGTACCTTTTGAATACATTAAAATAAATGCCTACTACAATATCAATAATATA 420
              ||  || |||||  ||||| |||| || || |||||| |||| || ||| |||||||
Db    975 CCCCACATTTGTTGCTTTAGAGTCATTAATCAACGAATCCATTAGCTTGATGCACTAA 916

Qy    421 TTCAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGCATCCCATGGG 480
              |||| ||| ||||||   ||| ||||| |||||||| || |||| ||| || |||||
Db    915 TTGAAAACGATGATCTAACCCTTTGAAAATGACGAAGTGGGGTAGGAATTGAATCTAAATT 856

Qy    481 AAGATTGGCACTGCTGCTAAGGCAAGTGCAGTAAGGCATTGAGTAGGTTATGCTTGCC 540
              |||||  ||| ||||  |||| ||||| |||||||| ||| || ||||||| |||||
Db    855 AATCCTATAGCTTGTGCCAATGCTGTTGCTGCCAAAAGTTCATATAATTATTATGGCGACC 796

Qy    541 TTTGATTTTTAATTGAATGGATTGGTAAAAGTTTGTTTTTTGATGATATAATGCCATACC 600
              |||||| ||| |||| || ||  ||| |||| ||||| |||| |||| | || ||||
Db    795 AACCAATGTAGCTTCTTCACAAGGTAAATCACTTCATCTTTACCATTAAAATATTTGCTT 736

Qy    601 ATCACCATCATCAAATAATAAAAATCTGC--CGTTTGGTGGCTAAGTAGCTATTTGTTG 659
              ||||||||||| |||||| |||||||  ||| |||| || |||| ||| | |||||
Db    735 GCCATTTCAGTTTTTAGCCAATAATCCCGACACTATTTTCCGCAAAAGAAAACGGTAGTTT 676

Qy    660 TGATAATGGTGTGATTGGGGGCAACGCTTGTCAGTGTCGTGTCAAGCATTGCTTGGCTAAAT 719
              ||||| | |||  ||| || ||| || | || |||||||| |||| ||||| |||||||
Db    675 CGCTTGATTTTCGTTTTCCCAAAAGTCAGCC------------------------ 644

Qy    720 TGGCCATCGTCTTGGGTAGATGATGGGCGGGTTGGCAATTTTCAAAAAATACCCAATTTTGTG 779
              |||||| || ||| ||| ||| || ||||  ||| ||| ||||   |||| |||||||
Db    643 ---TATCTTCATTGTTCAACACAACCTACTTTAGCATTTAAGTCACTTATGATAAAATGGGTAATTTTGCTT 587

Qy    780 CCAAATAATCCACCATCAAATAAATAAAATCTGC--CGTTTGGTGGCTAAGTAGCTATTTGTTG 839
              ||||||| ||||||| ||| ||||| || ||||| ||||| || |||| ||| | |||||
Db    735 GCCATTTCAGTTTTTAGCCAATAATCCATATAGCGGATCCATATAGCCTAGTCAGTTCAAGAGTAGTGTTT 676

Qy    840 TTGCCCCTTGAGCTGCCAAATTTGAGAATGCTCAAGCTGAAAACTTGACACGCTTCAGCA 899
              |||  |||||||||||| ||| |||| |||| ||||||| |||||||||||||||||||
Db    526 TCGCTGCCGCGACGCTTTTAGCTAYTAAGTTGTCTCAAGCTAAAAGCCTCTAGTA 467

Qy    900 CCACCAAATCCAATATTTGGATCGTTAATAATTCAAGTGCAGGGGTGCCAATAATTACCGC 959
              ||||||| || ||| ||||| | ||| || ||||| || |||| |||||||| |||||
Db    466 CATAAAGTTCACAAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAAATATTTCCGC 410

Qy    960 CGACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAAATGTCGTAACAGTGC 1019
              || || || ||||| || ||||  ||||||| | ||| |||| |||| | || |||||
Db    409 CCATACCAACTTTCACACCAGCAGCTTTCGCCATTTCATATAAAACTAAAAGTTAGCGGTAC 350

Qy   1020 TTTTTGGCATTGGAGCCTGATGATGGCGGATGAGTTGGTGTGTTGCTTGAGTTAGATCTC 1079
              ||||| || ||| || ||| ||| ||| || || ||||| | |||| ||||| || ||
Db    349 TTTTTACCCATTTGAACCTGTAATCCCCACAATTGGCTTTTGTCGCTGCGCGGCA------ 298

Qy   1080 GTGCTTTGAGTGTATCAATGAACAATTGACATCGCCTGATCACAGGAATGCCTTGAGCTT 1139
              ||||| ||| |||| |||| || ||||| || ||||| || | ||||| |||||||||
Db    297 ------GAATAATTCAATATGCGGCGATTACTTCCACTCCCGCTTTAA 257

Qy   1140 TGGCGGGCGACAATACTCGGCGTTCTTGGGGTTAATCGCCGATTGATGATCTGATCGG 1199
              |||| ||| ||||| || |||| ||||| |||||| |||| |||| ||| ||| ||
Db    256 GTGCGGGTTTGAAATTCCTCGGTGTTTTTACCCGGCGCTTAATAACAATCATATCGC 197

Qy   1200 CTGAGCTTAAGAGTTCACTGTCACAAAATGACAAATGAACGCCTGTTGGGCACTT 1259
              || || |||| || || ||||| ||||| |||| ||| | || || |||||||
Db    196 TTTCAAGTAACCATTCCTGATTTAAACTACCAGTATGAAGAGGGGATTATTTTGAGGAAGTT 137

Qy   1260 TATCTGCCAAGGTAGGAGGATTGGGATTACCCATCGATGATGAATGCGTAACTTTGTGACCGTGATTGA 1319
Qy   1260 TATCTGCCAAGGTAGGAGGATTGGGATTACCCATCGATGATGAATGCGTAACTTTGTGACCGTGATTGA 1319
```

```
PR   27-DEC-2002; 2002US-0436734P.
PR   27-DEC-2002; 2002US-0436804P.
PR   27-DEC-2002; 2002US-0436834P.
PR   27-DEC-2002; 2002US-0436842P.
PR   27-DEC-2002; 2002US-0436861P.
PR   27-DEC-2002; 2002US-0436885P.
PR   27-DEC-2002; 2002US-0436889P.
PR   27-DEC-2002; 2002US-0436893P.
PR   27-DEC-2002; 2002US-0436900P.
PR   27-DEC-2002; 2002US-0436947P.
PR   30-DEC-2002; 2002US-0436971P.
PR   30-DEC-2002; 2002US-0436987P.
PR   30-DEC-2002; 2002US-0437013P.
PR   30-DEC-2002; 2002US-0437038P.
PR   30-DEC-2002; 2002US-0437141P.
PR   31-DEC-2002; 2002US-0437281P.
PR   31-DEC-2002; 2002US-0437527P.
PR   31-DEC-2002; 2002US-0437620P.
PR   31-DEC-2002; 2002US-0437638P.
PR   02-APR-2003; 2003WO-CA000462.
PR   02-APR-2003; 2003WO-CA000464.
PR   04-APR-2003; 2003WO-CA000481.
PR   08-APR-2003; 2003WO-CA000485.
XX
DR   WPI; 2005-628189/64.
DR   P-PSDB; AEC10869.
XX
PI   Virag C;
PI   Ouyang H, Pinder B, Richards D, Tal M, Thalakada R, Vallee F;
PI   Kanagarajah D, Li Q, Mansoury K, Nethery-Brokx K, Ng I;
PI   Beattie B, Buzadzija K, Canadien V, Domagala M, Houston S;
PI   Edwards A, Dharamsi A, Veddadi M, Alam MZ, Arrowsmith C, Awrey DE;
XX
PA   (AFFI-) AFFINIUM PHARM INC.
XX
PT   New composition comprising purified polypeptides from bacteria (e.g.
PT   Escherichia coli), useful for diagnosing, preventing or treating
PT   microbial infections, or in pharmacogenomic or drug screening procedures.
XX
PS   Claim 65; SEQ ID NO 400; 667pp; English.
XX
CC   The invention relates to a composition (II) comprising purified
CC   polypeptides from bacteria. Also described: (1) a crystallized,
CC   recombinant polypeptide comprising an amino acid sequence of (I), where
CC   the polypeptide is in crystal form; (2) a crystallized complex comprising
CC   the crystallized, recombinant polypeptide and a co-factor or a small
CC   organic molecule, where the complex is in crystal form; and (3) a host
CC   cell comprising a nucleic acid encoding a polypeptide of (I), where a
CC   culture of the host cell produces at least about 1 mg of the polypeptide
CC   per liter of culture and the polypeptide is at least about one-third
CC   soluble as measured by gel electrophoresis. The composition and methods
CC   are useful for diagnosing, preventing or treating diseases, such as
CC   microbial infections. These may also be used in pharmacogenomic or drug
CC   screening procedures. The present sequence represents a Haemophilus
CC   influenzae UDP-N-acetylmuramoylalanine-D-glutamate ligase gene sequence,
CC   which is used in an example from the present invention.
XX
SQ   Sequence 1314 BP; 403 A; 214 C; 294 G; 403 T; 0 U; 0 Other;
```

```
Query Match            11.6%;  Score 160.4;  DB 14;  Length 1314;
Best Local Similarity  49.6%;  Pred. No. 2.5e-35;
Matches  683;  Conservative   0;  Mismatches 606;  Indels  89;  Gaps  7;
```

```
Db   1200 TGTTGGGCGTAAAAATTCTATCGCGTTGTTCCATTGTATCGAACAA------ATAACTTTG 1147

Qy    181 GAGTGAATTCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACTATTCCACCA   240
Db   1146 CGATGAAAATTTTGCAAGCAGCAGCGCACCATCTCGACCCAAAAACAATAACTATTCCACCA 1087

Qy    241 CTGCATAACATATGCTGTGACAGCTCACTCAAAAATCTTGACCTTTTACCAGACCAAAAT  300
Db   1086 TTGATTTAATTAATTCAGCTATTTTCTGAAAAAATCAGCGCCCCTTTTCCGTCTCCCTAGCCAA 1027

Qy    301 CAATGCCAGTGATGAATGCCATAAACCGTGCCAATGCTCAAGCCCATCAAGGGTTGA   360
Db   1026 CAAATGCAAATTTACCCTCAATATA--------AAGCCCAGCCAAATGCGAACAACTGTACT  976

Qy    361 GCCGATGTTTGTACCTTTTGAATCATTCAGTGTAAAAGTTGTTTTTTTGATGATAATGCCATACC 420
Db    975 CCCCACACATTTGTTGCTTGAGTCATTTAAATCACTTACCATTTAAATATGCTT  916

Qy    421 TTCACAGCGATGCGTGAGCCCCTTTGAAAAGTCTTAAGGGTATCAAGCATGGGATCCATGGG 480
Db    915 TTGAAAACGATGATCTAAACCCTTTGAAAATGACGAAGTGCGGTTACGGAATTGAATCTAAATT 856

Qy    481 AAGGATTGGGCAGCTGTTGTCCTAAGGCAAGTGCGAGATAAGGCATTGAGGTAGGTTATGCTTGCC 540
Db    855 AATACCTTATAGCTTGTTGCCAATGCTGGGAAAGTTCATATAATAATTATTGGGGAC  796

Qy    541 TTTGATTTTTTAATTGATTGGTTTAAAAGTTTTTTTTTGATGATATAATGCCATACC  600
Db    795 AACCAAATGTAGCTTCTTCTTCACAAGGTAAAAATCACTTCATCTTTACCATTAAATATTGCTT 736

Qy    601 ATCACCATCATCAAAATAAATAAAAATCTGC--CGTTTGGTGGCTTAAGTAAGCTATTGTTG 659
Db    735 GCCCATTTTCAGTTTTGTAGCCAATAATCCGCACTATTTTCCGCAGAAAAAGACGGTATGTTT 676

Qy    660 TGATAATGGGTGTTGAATTGGGGGCCAACGGCTTGTCAGTGTGGGTCAAGCAGCATTGCTTGGCTAAAT 719
Db    675 CGCTTGATTTCGTTTTCCCGAAAAGTCAGCC-------------------------------------- 644

Qy    720 TGGCATCGTCTTGGTAGATGATTGGCGGTTTGGCAATTTCAAAAATACGCAAATTTTTGTG 779
Db    643 ---TATCTTCATTGTTCAACACACTTACTTTAGCATTATGATTATGGATAAATGCGTAATTTTGCTT 587

Qy    780 CCAAAATAAATCCACCATGCCATCGTCCCGACTCAAGAGTGATCAGCGGAGAGATTTAAATTG 839
Db    586 GGCGATTAATTCTTCTAAATCATTATACGGATCCATAAGTGCTTCAGTCACGGTTCAAGACAG 527

Qy    840 TTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCA 899
Db    526 TCGCTCGCCGCAGCTTTTAAGCTATAAAGTGTCTCAAGCTGAAAAACTTAGACGCTCTAGTA 467

Qy    900 CCACCAAATCCATATTTGGATCGTTAATAAATTCAAGTGCAGGCGTGCCAATAATTACCGC 959
Db    466 CATAAAAGTTCACAAATCTTCATTC---AACAATGACAAAGCGGAATCCCAATAATTCCGC 410

Qy    960 CGACACCAACAAATCATGCCTGCATGTTTTCTGCCATCTCGCCTACTTAATGTGTAACAGTGC 1019
Db    409 CCATACCAACTTTCACACACCAGCTTTCGGCCATTTCATTAAAGTAGTTACCGTAC  350

Qy   1020 TTTTGGCATTTGAGCCTGTGATGGGCGATGATGGGTCGTTGTTGTTGTTGCTTGAGATCTC 1079
Db    349 TTTTACCATTTGAACCTGTAATCCCCACAATTGGCTTTGTGTCGCCGCGGCA------- 298

Qy   1080 GTGCTTTGAGTGTATCAATGAACAATGGACATCGCTTGATCACACGAGAAATGCCTTGAGCTT 1139
Db    297 ------------------------GAATAATTCAATATCGGCGGATTCATTCCACTCCGGTTTAA 257

Qy   1140 TGGCCGGCGACAATACTCGGGGTTCTTCGGGGTTAAATATCGGCGATTACTTCCACTCGATCGG 1199
Db    256 GTGCGCGGTTTGAATTTCTGGTGTTTTTACCGGAAGCCGTTAATAACAATCATATCGC  197

Qy   1200 CTGAGCTTAAGAGTTCACTCGTCCAAGCTGCCAAAAATGACAAATGAACGCCTGTGTGGGCAGTT 1259
Db    196 TTTCAAGTAACCATTCCTGATTTAAACTACCAGTATGAGAGGGGGATATTTGTGTAGGGGAAGTT 137
```

```
Qy      1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATTATAAATTGATTGATC   60
Db   1308 TAAACATTGAGCTAAAATGCGTAAATTCTTCGCCGGGCTTTTCAAAGAGAAGCAAAACTGATC 1249

Qy     61 CATGCTAGCACAAGCAGGGTGACAAAAAGCAAAGCCCTAGCGGTGCTTTGACTGGCTGCCAA  120
Db   1248 GAGACTTGCACAAGCAGGCGCACAATTAATTACCA----------------TATCTCCGCTTTGCAA 1201

Qy    121 TACTGCCCGATCAAGTAGTACCAACCTGATATATTTTTAGATGCCAAAGTGTCATCACCTTT  180
```

```
Qy  1260 TATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGA 1319
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   136 TATCAATACCAGTAGGATTTTTCGAGTATCAAATCACACAGAATATTAGCCTGTTGGGATA 77

Qy  1320 CCAAAAAATTCACAGCTGACAGTGCCCAGCTGTGCCCAAACCGACGACGGGCAATATTTTT 1377
         |||||| ||||||||||||  |||||   |||  || |  |  ||||||||||||||||||
Db    76 AGAGTAATCCACACAAGAAAGACCTGTTTTGCCAAGCCCGATGATGAGTAATATTTTT 19
```

RESULT 10
AEE31251/c

ID   AEE31251 standard; DNA; 2925 BP.
XX
AC   AEE31251;
XX
DT   09-FEB-2006 (first entry)
XX
DE   Haemophilus influenzae murD coding sequence.
XX
KW   vaccine; antibacterial; murD; coding sequence; ds.
XX
OS   Haemophilus influenzae.
XX
PN   US2005272089-A1.
XX
PD   08-DEC-2005.
XX
PF   01-AUG-2005; 2005US-00194246.
XX
PR   19-OCT-2001; 2001US-0345438P.
PR   21-OCT-2002; 2002US-00274586.
XX
PA   (MOTT/) MOTT J E.
PA   (TREP/) TREPOD C M.
PA   (ARVI/) ARVIDSON S.
XX
PI   Mott JE, Trepod CM, Arvidson S;
XX
DR   WPI; 2006-037204/04.
DR   P-PSDB; AEE31355.
XX
PT   Novel isolated polypeptide critical for survival of Haemophilus species,
PT   useful for reducing virulence of H. influenzae to be used as vaccine
PT   composition.
XX
PS   Example 3; SEQ ID NO 220; 158pp; English.
XX
CC   The invention relates to an isolated polypeptide (I) critical for the
CC   survival of Haemophilus sp. comprising an amino acid sequence having at
CC   least 95% structural similarity with an amino acid sequence chosen from
CC   AEE31317, AEE31330, AEE31332, AEE31335, AEE31353, AEE31354 and
CC   AEE31358. Also described: (1) an isolated polynucleotide (II) critical
CC   for the survival of Haemophilus sp. comprising a nucleotide sequence
CC   having at least 95% structural similarity with nucleotide sequence chosen
CC   from a coding sequence in AEE31062, AEE31130, AEE31140, AEE31155,
CC   AEE31211, AEE31241, AEE31246 and AEE31263 and AEE31231, and its
CC   complements, and identifying (M1) an agent that binds a polypeptide,
CC   encoding (I); (2) identifying (M1) an agent that binds a polypeptide,
CC   involves combining a polypeptide and an agent to form a mixture, where
CC   the polypeptide is encoded by a coding sequence comprising a nucleotide
CC   sequence chosen from AEE31062, AEE31130, AEE31140, AEE31155,
CC   AEE31211, AEE31241, AEE31246, AEE31263 and AEE31231, or has at least
CC   95% structural similarity with an amino acid sequence chosen from AEE31317, AEE31330,
CC   AEE31332, AEE31335, AEE31345, AEE31353, AEE31354 and AEE31358, and
CC   determining whether the agent binds the polypeptide; (3) decreasing the
CC   growth rate of a microbe, involves combining a microbe with an agent that
CC   binds to a polypeptide encoded by a coding sequence comprising a
CC   nucleotide sequence chosen from AEE31062, AEE31130, AEE31140, AEE31155,
CC   AEE31211, AEE31241, AEE31246, AEE31263 and AEE31231; (4) making (M2) an
CC   H. influenzae with reduced virulence, which involves altering a coding
CC   sequence in an H. influenzae to comprise a mutation, the non-mutagenized
CC   coding sequence comprising a nucleotide sequence chosen from AEE31062,
CC   AEE31130, AEE31140, AEE31155, AEE31211, AEE31241, AEE31246, AEE31263 and
CC   AEE31231, and determining if the H. influenzae comprising the mutation
CC   has reduced virulence compared to an H. influenzae that does not comprise
CC   the mutation; (5) an H. influenzae obtained by (M2); and (6) a vaccine
CC   composition comprising the polypeptide. (I) is useful for reducing the
CC   virulence of H. influenzae to be used as a vaccine composition, and for
CC   decreasing the growth rate of a microorganism. The present sequence
CC   represents a H. influenzae murD coding sequence, which is used in the
CC   exemplification of the present invention.
XX
SQ   Sequence 2925 BP; 796 A; 454 C; 640 G; 1035 T; 0 U; 0 Other;

Query Match   11.6%; Score 160.4; DB 15; Length 2925;
Best Local Similarity   49.6%;
Matches 683; Conservative   0; Mismatches 606; Indels 89; Gaps   7;

```
Qy     1 TAAAGGGTGAACCATCGTTAACAAACTTATCACCTGCTCATTGATGATTTTTAAAAATTGATC 60
                                                           
Db  2240 TAAACATTGAGCTAAAATGCGTAAAATTCTTCGCCGCGGCTTTTCAAAAGAGAAACTGATC 2181

Qy    61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTTACGCGGTGCTTTGGCTGCCCAA 120
         |||||||||||||||||||||||||||||||  |||||      |  ||| |||||||||||
Db  2180 GAGACTTGCACAAGCAGGCGACAATAATACCA----------TATCTCCGCTTTGCAA 2133

Qy   121 TACTGCCGATCAAGTGTACCAAACCTGATATATTTTTAGATGCCAAAATGTGCATCACCTTT 180
         |||||| |||||  |||||||||||||  |||||||||||||||  ||||||||||||||
Db  2132 TGTTGGGCGTAAAAATTCTATCGCTTGTTCCATTGTATCGAACAA----ATAACTTTG 2079

Qy   181 GAGTAAATCTTGCTCAATTATTCGGGGCATCTTGACCGGATTAAATACACACTACCACCA 240
         |||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db  2078 CGATGAAAAATTTTGCAAGCAGCGGCACCATTCTGACCAAAAATAACGTGTGG 2019

Qy   241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTTACCAAGCCCACCCAAAAT 300
         ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2018 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCGTCTCCGCCTAGCAA 1959

Qy   301 CAAATGCCAGTGATGATATGCCATAAACCGTGCCCAAGCCCATCAAATTGCTGCAACGGTTGA 360
         |||                  |                     || |||||||||||||
Db  1958 CAAATGCAATTACCCTCAATATA-----AAGCCCAGCCAATGCAGCGAGAACGTTACT 1908

Qy   361 GCCGATGTTGTTGTACCTTTTGAATCATTAAAATATGCCTTGCTTCAATATCAATAATA 420
         ||||||||||||  |||   |||||||||||||||||||  ||||||||||||||
Db  1907 CCCCACAATTGTTCGTTTGCTTTAGAGTCATTAATCCAACGAAATGCCATTAGCTTGATGACTAA 1848

Qy   421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGTATCAAGCATGGCATCCCATGGG 480
         ||||||  |   |||||||||||| ||||||||||     ||||  |||||||||||
Db  1847 TTGAAAACGATGATCTAACCCTTTAAATGAGTGCGGTACGGAATTGAATCTAAATT 1788

Qy   481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
         |||||||||||        ||||||||||||||||||||||||||||||||||||||
Db  1787 AATACCTATAGCCTTGTGCCAATGCGTTGCTGCCAAAAGTTCATTATAAGATTTATGGGACC 1728

Qy   541 TTTGATTTTAATTGATGGGATTGGTTAAAGTTTGTTTTTTTTGATGATAATGCCATACC 600
         |||  ||    |  |||||||||  ||||||| |||||||||||||||||||||||||
Db  1727 AACCAATGTAGCTTCTTCACAAGGTAAAATCATCTTTTACCATTAAAATATGCTT 1668

Qy   601 ATCACCATCATCAAAATAAATAAAAATCTGC-CGTTTGGTGGCTAAGTAAGTATTTGTTG 659
         |||||||||||||||||   |||||||||||| ||||||||||||||||||||
Db  1667 GCCATTTCAGTTTTTAGCCAATAATCCGCACTATTTTCCGCAAAAGAAACGGTAGGTTT 1608

Qy   660 TGATAATGGGTGTGATTGGGGCCAAGCCGCTTGTCAGTGTGGTCAAGCAATTGCTGCTAAAT 719
         |||||  ||||  |||||||||||||||||||||||||||||||||||||||
Db  1607 CGCTTGATTTCGTTGGTTTCCCAAAAGTCAGC---------------------1576

Qy   720 TGGCATCGTCTTGGTAGATGATGGCCGGTTGGCAATTTTCAAAAATACGCAATTTTTGTG 779
         ||||||||||  |||||  |||||||||| |||||   |  |||||||||  |||||
Db  1575 ---TATCTTCATTGTTCAAACACTACTTTAGCATTATGATAAATGCGTAATTTTGCTT 1519

Qy   780 CCAAATAATCCACCATGCCATCCGTGCCGATCAAGATGAATCAGCGGAGAGATTTAAAATTG 839
         |||       |||||||||||||||||||||||||||||||||||||||||
Db  1518 GGCGAAAAATCTTCTAAATACCATATAGCGATCCATATGATCTTCAGTCACGTTCAAGACAG 1459

Qy   840 TTGCCCCTTGAGCTGCCAAATTTGAGGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA 899
```

RESULT 11

AAT42063 11/c

Continuation (12 of 19) of AAT42063 from base 1100001

WP Sequence split into 19 fragments   LOCUS AAT42063 Accession Aat42063 (Haemophilus influenzae complete g

WP Fragment Name	Begin	End
AAT42063_00	1	110000
AAT42063_01	100001	210000
AAT42063_02	200001	310000
AAT42063_03	300001	410000
AAT42063_04	400001	510000
AAT42063_05	500001	610000
AAT42063_06	600001	710000
AAT42063_07	700001	810000
AAT42063_08	800001	910000
AAT42063_09	900001	1010000
AAT42063_10	1000001	1110000
AAT42063_11	1100001	1210000
AAT42063_12	1200001	1310000
AAT42063_13	1300001	1410000
AAT42063_14	1400001	1510000
AAT42063_15	1500001	1610000
AAT42063_16	1600001	1710000
AAT42063_17	1700001	1810000
AAT42063_18	1800001	1830121

Query Match   11.6%;   Score 160.4;   DB 2;   Length 110000;
Best Local Similarity   49.6%;   Pred. No. 2.1e-34;
Matches   683;   Conservative   0;   Mismatches   606;   Indels   89;   Gaps   7;

(DNA sequence alignment data — multiple rows of Qy and Db aligned sequences with position numbers)

RESULT 12
AAT42063_12/c
Continuation (13 of 19) of AAT42063 from base 1200001
WP Sequence split into 19 fragments LOCUS AAT42063 Accession Aat42063 (Haemophilus influenzae complete g

Fragment Name	Begin	End
AAT42063_00	1	110000
AAT42063_01	100001	210000
AAT42063_02	200001	310000
AAT42063_03	300001	410000
AAT42063_04	400001	510000
AAT42063_05	500001	610000
AAT42063_06	600001	710000
AAT42063_07	700001	810000
AAT42063_08	800001	910000
AAT42063_09	900001	1010000
AAT42063_10	1000001	1100000
AAT42063_11	1100001	1210000
AAT42063_12	1200001	1310000
AAT42063_13	1300001	1410000
AAT42063_14	1400001	1510000
AAT42063_15	1500001	1610000
AAT42063_16	1600001	1710000
AAT42063_17	1700001	1810000
AAT42063_18	1800001	1830121

Query Match   11.6%; Score 160.4; DB 2; Length 110000;
Best Local Similarity   49.6%;   Pred. No. 2.1e-34;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

RESULT 13
AEB39175_30
Continuation (31 of 35) of AEB39175 from base 3000001 (L. pneumophila DNA SEQ ID NO 3507.
WP Sequence split into 35 fragments LOCUS AEB39175 Accession Aeb39175

Fragment Name	Begin	End
AEB39175_00	1	110000
AEB39175_01	100001	210000
AEB39175_02	200001	310000
AEB39175_03	300001	410000
AEB39175_04	400001	510000
AEB39175_05	500001	610000
AEB39175_06	600001	710000

[Remainder of page consists of nucleotide sequence alignment lines (Qy/Db and Db/WP pairs with match bars) that are not transcribed base-by-base.]

```
WP  AEB39175_07           700000                                            810000
WP  AEB39175_08           800000                                            910000
WP  AEB39175_09           900000                                           1010000
WP  AEB39175_10          1000000                                           1110000
WP  AEB39175_11          1100000                                           1210000
WP  AEB39175_12          1200000                                           1310000
WP  AEB39175_13          1300000                                           1410000
WP  AEB39175_14          1400000                                           1510000
WP  AEB39175_15          1500000                                           1610000
WP  AEB39175_16          1600000                                           1710000
WP  AEB39175_17          1700000                                           1810000
WP  AEB39175_18          1800000                                           1910000
WP  AEB39175_19          1900000                                           2010000
WP  AEB39175_20          2000000                                           2110000
WP  AEB39175_21          2100000                                           2210000
WP  AEB39175_22          2200000                                           2310000
WP  AEB39175_23          2300000                                           2410000
WP  AEB39175_24          2400000                                           2510000
WP  AEB39175_25          2500000                                           2610000
WP  AEB39175_26          2600000                                           2710000
WP  AEB39175_27          2700000                                           2810000
WP  AEB39175_28          2800000                                           2910000
WP  AEB39175_29          2900000                                           3010000
WP  AEB39175_30          3000000                                           3110000
WP  AEB39175_31          3100000                                           3210000
WP  AEB39175_32          3200000                                           3310000
WP  AEB39175_33          3300000                                           3410000
WP  AEB39175_34          3400001                                           3503610
```

Query Match   11.4%;   Score 157.6;   DB 14;   Length 110000;
Best Local Similarity   51.9%;   Pred. No. 1.4e-33;
Matches 467; Conservative 0; Mismatches 409; Indels 24; Gaps 4;

```
Qy   170  GCATCACCCTTTGAGTAAAATCTTGCTCAATTATTCGGGGCATCTTGACCGATTAAATACACA  229
Db 39768  GGAACTACTTTGGCTAAAGCCGATTCAAATCTTGTCAGCATCTTCACCAAATCAAAAACAATG 39827

Qy   230  CTATCCACCCACTGCATAACATATCCTGACGAGCTCACTAAAATCTTGACCTTTACCAAGC  289
Db 39828  GAACGAACAAAATTCTGAAAACTGGGCTGGGGTCTAAATCTTGAAAAATCCGCCCCTTTTCCCTGT 39887

Qy   290  CCACCCAAAATCAATGCCAGTGATGATATATGCCATAAAACCGTGCCAAGCCCATCAATTGCT  349
Db 39888  CCACCCGCAATAAGAACAATTTTCCCTTGCATAGAACC---TCCTATTCCAATTAAATAGCA 39944

Qy   350  GCAATGGGTTGAGCCGATTGTTTTGTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATA  409
Db 39945  GAGATTGTAGCACACAATATTTGTTCCTTTCGAATCATTAATCCAACCTACTCCATCCACT 40004

Qy   410  TCATCAAATATATTCACAGCCGGATCGGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATG  469
Db 40005  TCTCTTACCCACTGGCATCGATGTGGCACGTCCAGGGAAGGTTTTTAAAACATTAGGAATG 40064

Qy   470  GCATCCATGGGAAGATTGGCAGCTGCTCGCAGTAAAGGCATTGAGTAGG  529
Db 40065  TGTTGCATAGAGATTCCCGCTGCTTCGGGCTAAAGCACAAGCGGCATTCATCCAA 40124

Qy   530  TTATGCTTGCCTTTGTATTTTAATTGATGGATTGGTAAAAGTTGTTTTTTTGATGATAT  589
Db 40125  TTATGCACCCCTTTTATTAAAATCGATTCCACAGGCAAAAGACGCTCCATTCCTTTCGCC 40184

Qy   590  AATCCATACCACCATCACTCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAG  649
Db 40185  AAATAAGTAGTATTCTCTGTTCAATCACAGCCCCCAAATTCCCATT------GAAG 40233

Qy   650  CTATTGTGTGATAATGGTGATTGGGGCAACAGCTTGTCAGTGTGGTCAAGCGATTGC  709
Db 40234  GTGCGTCTTTGCCAAAAGAGATGCATTTTATATCTGTTTGACAGGATTGGTGAGGAACCGG 40293

Qy   710  TTGGCTAAAATTGGGCATCGTCTTGGTAGAGATGATGGCGGGTTTGGGAATTTTCAAAAAATACGC  769
Db 40294  T-------ATAGACGCTTCTCGATTAAAATAAAAACAGCCTTGGCACCAGCGATAAAATCCTT 40346
```

RESULT 14
AEB42401_28
Continuation (29 of 34) of AEB42401 from base 2800001 (L. pneumophila DNA SEQ ID NO 6733
Sequence split into 34 fragments   LOCUS AEB42401 Accession Aeb42401

```
WP  Fragment Name                 Begin      End
WP  AEB42401_00          110000                                             110000
WP  AEB42401_01          100001                                             110000
WP  AEB42401_02          200001                                             310000
WP  AEB42401_03          300001                                             410000
WP  AEB42401_04          400001                                             510000
WP  AEB42401_05          500001                                             610000
WP  AEB42401_06          600001                                             710000
WP  AEB42401_07          700001                                             810000
WP  AEB42401_08          800001                                             910000
WP  AEB42401_09          900001                                            1010000
WP  AEB42401_10         1000001                                            1110000
WP  AEB42401_11         1100001                                            1210000
WP  AEB42401_12         1200001                                            1310000
WP  AEB42401_13         1300001                                            1410000
WP  AEB42401_14         1400001                                            1510000
WP  AEB42401_15         1500001                                            1610000
WP  AEB42401_16         1600001                                            1710000
WP  AEB42401_17         1700001                                            1810000
WP  AEB42401_18         1800001                                            1910000
WP  AEB42401_19         1900001                                            2010000
WP  AEB42401_20         2000001                                            2110000
WP  AEB42401_21         2100001                                            2210000
WP  AEB42401_22         2200001                                            2310000
WP  AEB42401_23         2300001                                            2410000
WP  AEB42401_24         2400001                                            2510000
WP  AEB42401_25         2500001                                            2610000
WP  AEB42401_26         2600001                                            2710000
WP  AEB42401_27         2700001                                            2810000
WP  AEB42401_28         2800001                                            2910000
WP  AEB42401_29         2900001                                            3010000
WP  AEB42401_30         3000001                                            3110000
WP  AEB42401_31         3100001                                            3210000
WP  AEB42401_32         3200001                                            3310000
WP  AEB42401_33         3300001                                            3345687
```

Query Match   11.4%;   Score 157.6;   DB 14;   Length 110000;
Best Local Similarity   51.9%;   Pred. No. 1.4e-33;
Matches 467; Conservative 0; Mismatches 409; Indels 24; Gaps 4;

```
Qy   770  AATTTTTTGTGCCAAATAATTCACCATCGTCGCGATCAAGATGATCAGCGGAGAGA  829
Db 40347  TGTTTTGCTTCGCGTATAGGCTTCCATAGTGTGATGGCTTCAAATGATCAGGCGGTTACA 40406

Qy   830  TTTAAAATTGTTGCCCTTGAGCTGCCAAATTTGAGATATGCTTCAAGCTGAAAACTTGAC  889
Db 40407  TTAAGAATCGTCGCAACCACTGGGCACAAAGAATAAGTTAAAATCCAAATTCAAAACTGGAT 40466

Qy   890  AGCTCAAGCGACCACCACCAAAATC---CATATTTGGATCGTTAATTAATTCAAGTGCAGCGGGTG  946
Db 40467  AACTCCCAAATAACCCACAAAATCGTAGTGAATGGTTCATCATCTAAGACACAGGAGTG 40526

Qy   947  CCAAATATTACCGCCCGACACACAACATCATGCCTGCTGCAGTGTTTTTGCCATCTCCCTTAAT  1006
Db 40527  CCGATAATTTCCGGCTACAGCTTACGCGCCGAAAACCCGGCCGCCTTTGCCACCAAA 40586

Qy  1007  GTCGTTAACAGTGCCTTTTGAGCCTTTGTTGAGCCTCGTGATGGCGATGAGTTGGTGGTGTTTGTTGCT  1066
Db 40587  GTCGTTACAGTGGATTGCCATTAGTTCCCGGTTATGGCAATGACAGGGGGGGGCTAAHCTCT 40646
```

Qy   170 GCATCACCCTTTGAGTAAAATCTTGCTCAATTATTCGGGGCATCTTGACCGATTAAATACACA 229
Db 96841 GGAACTACTTTGGCTAAAGCCGATTCAAATCTTGTCAGCATCTTCACCAAATCAAAAACAATG 96900

Qy   230 CTATCCACCCACTGCATAACATATCCTGACGAGCTCACTAAAATCTTGACCTTTACCAAGC 289
Db 96901 GAACGAACAAAATTCTGAAAACTGGGCTGGGGTCTAAATCTTGAAAAATCCGCCCCTTTTCCCTGT 96960

```
Qy    290  CCACCCAAAATCAATGCCAGTGATGATGATATGCCATAAAACGGTGCCAAGCCATCAAATTGCT 349
           ||||                                                      ||||||
Db  96961  CCACCCGCAATAAGAACAATTTCCCCTTGGCATAGAGAACC---TCCTATTCCATTAATAGCA 97017

Qy    350  GCAATGGTTGAGCCGATGTTTGTACCTTTTTGAATCATTAAAAATAATGCCTTGCTATCAATA 409
           |||                                                    |||
Db  97018  GAGATTGTAGCACCAAATATTTGTTTCCTTTCGAATCATTAATCCAACCTACTCCATTCCACT 97077

Qy    410  TCATCAATATATTCACAGCGGATCGGTAGCCCTTTGAAAGTCTTAAGGGTATTCAAGCATG 469
           |||||                                           ||||
Db  97078  TCTCTTACCCACTCGACATCGGATGTGGCAGTCCAGGGAAGGTTTTTAAAACATTAAGAATG 97137

Qy    470  GCATCCCATGGGAAGATTGGCAGCGTGTGCCTAAGGCAAGTGCCAGATAAGGCATTGAGTAGG 529
           ||||                                                  ||||
Db  97138  TGTTGCCATAGAAAATTCCGCGTGCTTCGGCCTAAAGCACAAGCCAGCAGCATTCATCCAA 97197

Qy    530  TTATGCTTGCCTTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATAT 589
           ||||                                          ||||
Db  97198  TTATGCACCCCCTTTTTTATTAAAAATCGATTCCACTGCCAAAAAGACCGCTTCCTTTTCGCC 97257

Qy    590  AATGCCATACCATCACCCATCATCAAATATAAAATAAATCTGCCGTTTGGTGGCTAAGGTAAG 649
           |||                                                    ||||
Db  97258  AAATAAAATAGTATTCTTGTTTCAATCAGCCCCCAAATTTCCCATT-------GAAG 97306

Qy    650  CTATTTGTTTGTGATAAATGGTGTGATTTGGGGGCACACGGCTTGTGTCACTGTGTGCTCAAGCCATTGC 709
           ||||                                                    ||||
Db  97307  GCGGCTCTTTGCCAAAAGAGATGCAATTTTATATCTGCCTTGACACAGGATTGGTCAGGA---- 97362

Qy    710  TTGGCTAAAATTGGCATCGTCTTTGGGTAGAGATCGATGATGGCGGTTTGGCAATTTCAAAAATACGC 769
           ||||                                                    ||||
Db  97363  ---ACAGTGTAGACGGTCTTCACGATTAAAATAAAACAGCCTTTGGGCACCTTTGGCACCTTTCCTT 97419

Qy    770  AATTTTTGTGCCAAATAATCCACCATGCCATCGTCGCCGATCAAGATGATCAGCGGAGAGA 829
           ||||                                                    ||||
Db  97420  TGTTTGCTTGCGTTATAGGCTTCCATAGTGTGATGCCTGTCCATAGTTGACTGAAAACTTAC 97479

Qy    830  TTTAAAATTGTTGCCCCTGAGCTCCCAAATTTGAGATATGCTCAAAGCTGATGAAAACTTGAC 889
           ||||                                                    ||||
Db  97480  TTAAGAATCGTCGACAAGAATAAGTTAAATCCAGTTGAAAACTGGAT 97539

Qy    890  AGCTCAAGCACACCAACCAAAATC----CATATTTTGGATCGTTAATAATTCAAGTGTGCAGGCGTG 946
           ||||                                                    ||||
Db  97540  AACTCCAAATACCCACAAATCGTAGTGTTCATCATCATCTAAGACACAGGAGTG 97599

Qy    947  CCAATATTACCGCGCGACCACAACAATCATGCCTGCATGTTTGCCATCTCGCCTACTAAT 1006
           ||||                                                    ||||
Db  97600  CCGATATTTCCGGCTACACGAAACCCGGCGCCCTTAGCCATTTCTCCACCAA 97659

Qy   1007  GTCGTAACAGTGCTTTTGGGCATTTTGGATGGCGATGATTGGTGTTTTGCT 1066
           ||||                                                    ||||
Db  97660  GTCGTTACAGTGTGGATTTACCATTAGTTCCCGTTATAGCAATGACAGGGGGGCTAATCTCT 97719
```

PT   New polynucleotides of nontypeable strain of Haemophilus influenzae,
PT   useful for treating or preventing NTHi bacterial infections of the middle
PT   ear and/or nasopharynx.
XX
PS   Example 1; SEQ ID NO 569; 88bp; English.
XX
CC   The invention comprises nucleotide sequences (genes) from the genome of a
CC   nontypeable strain of Haemophilus influenzae (NTHi). The NTHi DNA
CC   sequences of the invention are useful for treating or preventing NTHi
CC   bacterial infections of the middle ear and/or nasopharynx. The present
CC   nucleic acid represents an NTHi contig sequence of the invention.
XX
SQ   Sequence 14324 BP; 4572 A; 3171 C; 2271 G; 4310 T; 0 U; 0 Other;

Query Match   11.2%; Score 154.4; DB 13; Length 14324;
Best Local Similarity 49.0%; Pred. No. 4.3e-33;
Matches 591; Conservative 0; Mismatches 546; Indels 69; Gaps 4;

```
Qy    172  ATCACCTTTGAGTAAATCTGCCTCAATTATCGGGGGCATCTTGACCGATTAAATACACACT 231
           |||                                                    ||||
Db   5280  ATTAGCTTTGTGATGATGAAAATTTTGCAAGCTGCACACCATCTCGACCAAAACAATAACAAAT 5339

Qy    232  ATCCACCCACTGCATAACATATGCTGACACGTCTCACTAAAAAATCTTACCTTTACCAAGCCC 291
           |||                                                    ||||
Db   5340  AATGTGTGGTTGATTAAATTAAATTCAGCTAATTCTGAAAAAATCAGCCGCCCTTTTCCGTCTCC 5399

Qy    292  ACCCAAAATCAAATGCCAGTGATGATATGCCATAAAACCGTGGCCAAGCCATCAATTGCTGC 351
           |||                                                    ||||
Db   5400  GCCTAGCCAACAAAATGCAAT-------TTACCCTCAATAATAAAAGCTAATGCAGC 5450

Qy    352  AATGGTTGAGCCGATGTTTGTACCTTTTTGAATCATTAAAATATGCCTTGCTATCAATATC 411
           |||                                                    ||||
Db   5451  AACGTACTCCCCCACCATTTGTTTGCTTTAGTGTCATTAATCCAACCCACCATTAGCTTG 5510

Qy    412  ATCAATATATTCACAGCGATCGGTAGCCCTTTGAAAGTCTTAAGGGTATTCAAGCATGGC 471
           |||                                                    ||||
Db   5511  ATGCACTAATTGAAAAACGATGATCTAACCCTTTGAAATGACGAAGTGCGGTACGAATTGA 5570

Qy    472  ATCCATGGGAAGATTGGCAGCGTGCCTAAGGCAAGTGCGAGATAAGGCATTGAGTAGGTT 531
           |||                                                    ||||
Db   5571  AGTCAAATAATAACCGACACGCTTGTGTCGCCAAAAGTGTTCATATAAATT 5630

Qy    532  ATGCTTGCCTTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAA 591
           |||                                                    ||||
Db   5631  ATGGCGACGACCGACTAAATGTTACTTCCTTCACAAGGCAAAATTACTTCATCTTTTACCATTAA 5690

Qy    592  TGCCATACCATCACCATCATCAAAATAAATAAATCTGCCGTTTGGTGGCTAAGTAAGCT 651
           |||                                                    ||||
Db   5691  GTATTGCTTGCCCATTTTCAGTTTTTAGCCAATAATCCGACACTAGTTTCCGGCAAAAGAAAC 5750

Qy    652  ATTTGTTGTGATAAATGGTGTGATTTGGGGGCAAACGCCTTGTCAGTGTGGTCAAGCATTGCTT 711
           |||                                                    ||||
Db   5751  GGTTTGTTCTCGCTTGATTTTCACCTTCACCAAAAAGTCAGC-----------  5790

Qy    712  GGCTAAAATTGGCATCGTCTTGGGTAGGATGATGGCGGTTTGGCAATTTCAAAAATAGCAA 771
           |||                                                    ||||
Db   5791  ------TTATCCTCATTATTCAAAAACTCCGAGACCTCAGCATTTGATTAATTATCCGGTAA 5840

Qy    772  TTTTTTCTGTGCCAAATAATCCACCATGCCATGGTCGCCGAGATTCAGTGGGGAGAGATT 831
           |||                                                    ||||
Db   5841  TTTTGCTTGGGGATTAATTTCTTAAAATCCATATATAGCGGATCCATATGCAGCTTCAACGTT 5900

Qy    832  TAAAAATTGTTGCCCCTTGAGCTGCCAAAATTTGAGATATGCTCAAGCTTGAAAACTTGACAG 891
           |||                                                    ||||
Db   5901  ...
```

```
Db  5901 CAAAACAGTCGCAGCCGCGAGCTTTTAAGCTATATAAGTTGTCTCAAGCTGAAAACTAGAAG 5960

Qy   892 CTCAAGCACCACCAAATCCATATATTTTGGGATCGTTAATAATTCAAGTGCAGGGGTGCCAAT 951
         |||  ||| |||||||
Db  5961 CTCTAGTACATAAAAGTTCACAAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAAT 6017

Qy   952 ATTACCGCGACACCAACAAATCATGTCCTGCATGTTTTTGCCATCTCGCCTACTAATGTCGT 1011
         |||   |||||||      ||||||
Db  6018 ATTTCCGCCCATACCAACTTTCACGGCCAGCAGCTTTCGCCCATTTCATAAACTAAAGTAGT 6077

Qy  1012 AACAGTGCTTTTGGCATTTGAGCCCTGTGATGGGCGATGAATTGGGTGTGTTTGTTGCTTGAGT 1071
         |||||  ||||||||    ||||||
Db  6078 AACAGTACTTTTTCCATTTGAACCTGTAATCCCCACAATGGCTTTGTCGGCTGCGCG--- 6134

Qy  1072 TAGATCTCGTGCTTTGAGTGTATCAAATGAACAATTGAACATCGCTGATCACAGGAATGCC 1131
                                   |||||||||||| |||  |||||||
Db  6135 ----------------GCAGAATAAATTCAAATATCGCCGGATTACTTCCACTCC 6170

Qy  1132 TTGAGCTTTGGCGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTTGATGATGAT 1191
         |||     ||||||
Db  6171 CGCTTTAAGTGCGGGGTTTGAATTTCTGGTGTTTTTTACCGCAAAGCCCTGGGCTAATAACAAT 6230

Qy  1192 CTGATCGGGCTGAGCTTAAGAGTTCACTGTCTCCAAGCTGCCAAAAATGACAAATGAACGCCTGT 1251
         |||||
Db  6231 CATATCGCTTTCAAGTAACCATTCCTGATTTAAAACTACCAGTATGGAAGGAGGATATTTTTG 6290

Qy  1252 GGGCAGTTTATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTTGACC 1311
         |||||
Db  6291 AGGAAGTTTATCAATACCAGTAGGCTTTTTCGGGGTATCAATCACACGAATAATTAGCCTG 6350

Qy  1312 GTGATTGACCAAAAAATTCACAGCTTGACAGTCCGACAGAGCTGCCCAAAACCGACGACGGCATA 1371
         |||||
Db  6351 TTGGGATAAGAGATAATCCACACAAGAAAGACACCTCTTTTGCCCAAGCCCAATGATCGTAAT 6410

Qy  1372 TTTTTTT 1377
         ||||||
Db  6411 ATTTTTT 6416
```

Search completed: May 18, 2006, 13:52:28
Job time : 955 secs

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Biocceleration Ltd.

Run on: May 18, 2006, 08:54:03 ; Search time 6244 Seconds
         (without alignments)
         12358.853 Million cell updates/sec

Title: US-10-672-787-35_COPY_11357_12736
Perfect score: 1380
Sequence: 1 taaagcgtgaaccatgttaa.........acgacgggcatattttttga 1380

Scoring table: IDENTITY NUC
               Gapop 10_0 , Gapext 1.0

Searched: 48236798 seqs, 2795965780 residues

Total number of hits satisfying chosen parameters: 96455510

Minimum DB seq length: 22
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database : EST:*
    1: gb_est1:*
    2: gb_est3:*
    3: gb_est4:*
    4: gb_est5:*
    5: gb_est6:*
    6: gb_htc:*
    7: gb_est7:*
    8: gb_est8:*
    9: gb_est9:*
   10: gb_gss1:*
   11: gb_gss2:*
   12: gb_gss3:*
   13: gb_gss3:*
   14: gb_gss4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	98.2	7.1	817	13	CZ543630	CZ543630 SRAA-aads
c	2	92.8	6.7	709	13	CF306940	CF306940 HDA1-05-
c	3	92	6.7	892	12	BZ579116	BZ579116 msh2_6205
c	4	80.8	5.9	916	14	DU735587	DU735587 APKI3331.
c	5	80.2	5.8	1025	14	DU773354	DU773354 APKG3419.
c	6	77.2	5.6	555	14	AG943479	AG943479 Drosophil
c	7	76.4	5.5	977	14	DU748329	DU748329 Drosophil
c	8	69.8	5.1	1445	12	BZ568959	BZ568959 pacs2-164
c	9	62.2	4.5	623	14	AG952088	AG952088 Drosophil
c	10	59.6	4.3	928	14	DU769542	DU769542 Drosophil
c	11	56.8	4.1	628	13	CL691264	CL691264 PRI01561a
c	12	55	4.0	474	9	DR032465	DR032465 bda020270
c	13	54.2	3.9	1024	14	DU732353	DU732353 APKI1690.
c	14	52.6	3.8	1171	13	BZ559806	BZ559806 pacs2-164
c	15	50.2	3.6	838	13	CZ545153	CZ545153 SRAA-aad6
c	16	49.8	3.6	572	11	BH795426	BH795426 BMBAC376G
c	17	49.4	3.6	452	12	BE184674	BE184674 PM4-HT068
c	18	48.6	3.5	958	14	DU771470	DU771470 APKG2473.
c	19	47.6	3.4	1101	14	CNS00L00	AL068607 Drosophil
c	20	47.2	3.4	1101	14	CNS002MK	AL097862 Drosophil
c	21	46	3.3	823	10	DR505933	WS02715.B
c	22	45.4	3.3	856	14	AJ863943	Ralstonia
c	23	45	3.3	590	1	AU287701	AU287701
c	24	45	3.3	757	10	DR473894	WS00960.B
c	25	45	3.3	907	9	DN922769	43863.2.C
c	26	44.4	3.2	442	10	DR555020	WS02632.C
c	27	44.4	3.2	490	8	CO253110	WS00818.B
c	28	44.4	3.2	794	10	DR473810	WS00960.B
c	29	44	3.2	918	14	DU724406	APKG2941.
c	30	43.6	3.2	692	11	BH983612	ode20g11.
c	31	43	3.1	1035	14	CNS002D5	AL097523 Drosophil
c	32	42.8	3.1	1101	14	CNS0100X	AL098379 Drosophil
c	33	42.4	3.1	934	14	DU791861	APKH4300.
c	34	42.4	3.1	1101	14	CNS0039G	AL063921 Drosophil
c	35	42.2	3.1	747	13	CZ668937	OM_Ba022
c	36	42.2	3.1	915	11	CNS05L54	AZ543308 ENTGK93TR
c	37	42	3.0	941	14	CNS07ANB	AZ543308 Tetraodon
c	38	42	3.0	1022	14	AU287702	AL342337 Tetraodon
c	39	41.8	3.0	517	1	CNS000D1	AL065414 Drosophil
c	40	41.8	3.0	616	11	BH383040	AG-ND-127
c	41	41.6	3.0	1101	14	BH383040	AG-ND-127
c	42	41.6	3.0	714	11	BH374901	AG-ND-127
c	43	41.6	3.0	747	4	CA920035	EST637753
c	44	41.4	3.0	639	9	CX639436	UCRPT02.4
c	45	41.4	3.0	760	11	BH367641	AG-ND-I27

ALIGNMENTS

RESULT 1
CZ543630/c
LOCUS   CZ543630   817 bp   DNA   linear   GSS 13-MAY-2005
DEFINITION SRAA-aads1b03.g1 Strongyloides ratti genomic whole genome shotgun
            (SRAAGSS 004) Strongyloides ratti genomic, genomic survey sequence.
ACCESSION   CZ543630
VERSION   CZ543630.1 GI:64673383
KEYWORDS   GSS.
SOURCE   Strongyloides ratti
  ORGANISM Strongyloides ratti
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Panagrolaimoidea; Strongyloididae; Strongyloides.
REFERENCE   1 (bases 1 to 817)
  AUTHORS   Mitreva,M., McCatter,J.P., Thompson,F., Viney,M., Pape,D.,
            Ritter,E., Martin,J., Wylie,T., Dante,M., Waterston,R.H.,
            Clifton,S.W. and Wilson,R.
  TITLE   Genome Survey sequences from the rat parasitic nematode
            Strongyloides ratti
  JOURNAL   Unpublished (2005)
COMMENT   Contact: Mitreva M
            Washington University in St. Louis
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: nematode@watson.wustl.edu
            Genomic DNA was provided by Fiona Thompson
            (F.Thompson@bristol.ac.uk) and Mark Viney
            (Mark.Viney@bristol.ac.uk) at the University of Bristol, Bristol,
            UK.
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..817
                     /organism="Strongyloides ratti"
                     /mol_type="genomic DNA"
                     /strain="isofemale line ED321 heterogonic"
                     /db_xref="taxon:34506"
                     /dev_stage="infective larval stage (iL3)"
                     /lab_host="GS10"
                     /clone_lib="Strongyloides ratti whole genome shotgun
                     library (SRAAGSS 004)"
                     /note="Vector: pOTW13; Site_1: BstX1; Site_2: BstX1;

RESULT 2

ORIGIN

Strongyloides ratti genomic DNA was randomly sheared,
end-repaired and size fractioned to enrich for 2-4 kb
fragments. Genomic DNA was provided by Fiona Thompson
(F.Thompson@bristol.ac.uk) and Mark Viney
(Mark.Viney@bristol.ac.uk) at the University of Bristol,
Bristol, UK. Sequencing by Washington University Genome
Sequencing Center, St. Louis, MO."

Query Match   59.0%; Score 98.2; DB 13; Length 817;
Best Local Similarity   59.0%; Pred. No. 1.5e-16;
Matches 207; Conservative   0; Mismatches 138; Indels   6; Gaps   2;

Qy   722 GCATCGTCTTGGGTAGATGGAGATGGGCCGGTTGGCAATTTTCAAAAATACGCAATTTTTGTGCC 781
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   395 GCATCTTGTACGTTAACAAATGCAGTTTTTAGCATCGTTGTTATAAATTTTTTAGCTTCA 336

Qy   782 AAATAATCCACCATGCC----ATCGTGCCGATCAAGATGATCAGCGGAGAGAGATTTAAAATT 838
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   335 CGATACTCGCGCTAGACCTAATGGGTATCTATCCATATGATCTTCACTAATATTAAGACT 276

Qy   839 GTTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGC 898
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   275 GTCCGCGGGGGCGCCTTCTAAACTTGAAGTGGTTTCTAATTGGGAAGGCTGGAAAGTTCTAAA 216

Qy   899 ACCACCAAAATCCATATTTTGGATCGTTAATAATTCAAGTAGTCAGGGCGTTGCCAATATTACCG 958
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   215 ACGTATAGACG---ATGCCTGCATTGAGTAATGAAAACGGCTGCGTACCAAATATTGCCA 159

Qy   959 CCGACACCAACAATCATGCCTGCGATGTTTGCCATCTCGCCTACTAATGTCGTAACAGTG 1018
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   158 CCAACACCAACTGAAAATACCCGACGCTTTAGCCATTTCACCAACCAAATGACGTGACGGTA 99

Qy   1019 CTTTTTGGGCATTTGAGGCCGTGATGGGCGGATGAATTCGGCAATCGGCTTCACGA 1069
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db   98 CTTTTACCATTTGAACCTGTTAATGGCGACAAATCGGTCGCATCGGCTTCACGA 48

LOCUS   CF306940   709 bp   mRNA   linear   EST 15-AUG-2003
DEFINITION HDA1--05-E12.g1 OsHDAC1-overexpressing transgenic rice lambda phage
            cDNA library I (HDA1) Oryza sativa (japonica cultivar-group) cDNA
            clone HDA1--05-E12, mRNA sequence.
ACCESSION   CF306940
VERSION   CF306940.1 GI:33678701
KEYWORDS   EST.
SOURCE   Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; BEP
            clade; Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 709)
AUTHORS   Kim,J.S., Jun,K.M., Cheong,P.J., Kim,M.J., Lee,T.H., Shin,Y.C.,
            Song,S.I., Kim,J.K., Kim,Y.-K. and Nahm,B.H.
TITLE   Large-scale Sequencing Analysis of Rice ESTs
JOURNAL   Unpublished (2003)
COMMENT   Contact: Nahm B.H.
            Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
            of Bioscience and Bioinformatics, Myongji University
            Yongin, Kyeonggi, Korea
            Tel: 82 31 330 6193
            Fax: 82 31 321 6355
            Email: bhnahm@ggbio.com, bhnahm@bio.myongji.ac.kr.
FEATURES   Location/Qualifiers
     source   1..709
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /mol_type="mRNA"
                     /cultivar="Nackdong"
                     /db_xref="taxon:39947"
                     /clone="HDA1--05-E12"
                     /tissue_type="callus"
                     /dev_stage="proliferated callus on 2N6 media for 2 weeks"

RESULT 3

ORIGIN

"lab host="E.coli SOLR"
clone_lib="OsHDAC1-overexpressing transgenic rice lambda
phage cDNA library I (HDA1)"
note="Vector: pBluescript SK(+); Site 1: EcoRI; Site 2:
XhoI; Callus was treated with ABA(20um) for 1hour; cDNA
was inserted into lambda Uni-ZAP XR vector at 5' end with
EcoRI and 3' end with XhoI site. mRNA was derived from
rice Histone Deacetylase overexpression line."

Query Match   58.0%; Score 92.8; DB 5; Length 709;
Best Local Similarity   58.0%; Pred. No. 5.3e-15;
Matches 204; Conservative   0; Mismatches 142; Indels   6; Gaps   2;

Qy   721 GGCATCGTCTTGGTAGATGAGTAGATGGGCCGGTTGGCAATTTTCAAAAATACGCAATTTTTGTGC 780
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   449 GGCATCATCAGCATTAACCACGCAAAACTTCGGTTTTTCGGTAAATGCCGCAGTTTTGCTGC 390

Qy   781 CAAATAATCCACCATGCCA---TCGTGCCGATCAAGATGATCAGCGGAGAGAGATTTAAAAT 837
          ||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db   389 ACGATACIGTTGTAAACCAAACGGATAGCGGATCGCATCCAATATAGATCTTCAGTCAGGAAT 330

Qy   838 TGTTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAG 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   329 GGTCGCTGCTTGACCGCCCTGTAAGCTGGAGGTGGTTTCACCGGGAAGCCTGGACAGTTCAG 270

Qy   898 CACCACCAAATCCATATTTTGGATCGTTAATAATTCAAGTAGTCGTTAATATTCAGTATG 957
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   269 CACGTACAGTTCACACTCATCATCCAGTAGCATC---AACGCAGGCAGGCGTGCCAATAATTGCC 213

Qy   958 GCCGACACCAACAATCATGCCTGCGATGTTTGCCATCTCGCCTACTAATGTCGTAACAGT 1017
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   212 ACCCAACACCAACGTTAACCCCCGCCCGTTCGCGAACCCCCACTAGCGTGGTTGACCCGT 153

Qy   1018 GCTTTTTGGGCATTTGAGCCGTGTGATGGGCGGATGAATTGGGTGTGTTTGTTGCTTGA 1069
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   152 GCTTTTGCCGCTTAGAACCGTAATCGCCACAAATCGGTCGCTTGCTTGCTTGCTTCGGGA 101

LOCUS   BZ579116   892 bp   DNA   linear   GSS 17-DEC-2002
DEFINITION mbh2_6205.y2 msh Pseudomonas aeruginosa genomic clone mbh2_6205,
            genomic survey sequence.
ACCESSION   BZ579116
VERSION   BZ579116.1 GI:27214177
KEYWORDS   GSS.
SOURCE   Pseudomonas aeruginosa
ORGANISM   Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 892)
AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE   Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa
JOURNAL   J. Bacteriol. (2002) In press
COMMENT   Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 206221954
            Fax: 2066857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..892
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="MSH"
                     /db_xref="taxon:287"
                     /clone="mbh2_6205"
                     /clone_lib="msh"

ORIGIN   /note="Environmental isolate. Whole genomic shotgun
                library."

Qy   Best Local Similarity   56.3%;   Pred. No. 9.6e-15;
     Query Match   6.7%;   Score 92; DB 12; Length 892;
     Matches 192; Conservative   0; Mismatches 146; Indels   3; Gaps   1;

Qy   721 GGCATCGTCTTGGTAGATGAATGGGCGTTTGGCAATTTCAAAAATACGCAATTTTGTGCC 780
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db   256 GGCATCGGCGCGGATTCACCACGACCTGGCGCGGGCACCGGGAAGATCCGGTTGTGCTTGGCCAG 315

Qy   781 CAAATAATCCACCATCGTCCGATCAAGATGATCAGCCGAGAGAGATTAAAATTGT 840
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   316 GTGGTAGTCAGCCATCGCCCGTGTAGCGGATCCCATATGGGTCTTCGCTGACCTTCAGCACCGT 375

Qy   841 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAAGCTGAAAACTTGACAGCTCAAGCAC 900
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   376 CCGCACCTCGGCCGTTGAGCGGATCGCAGGTTTCCAGCTCGGAAGCTCGACAGCTCCAACAC 435

Qy   901 CACCAAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGGTTCCAATATTACCGCC 960
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   436 GTACAGCTCGATGTCGTCGGCC---AGCAGGTCGAGCGCGCCGGGGGTGCCCGAGGGTTGCCGGCC 492

Qy   961 GACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGCT 1020
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   493 GACGGCGACACGCTTGTCCGGCGGCCACCCGCCCATTTCGCCCACGGGTGGTCACGGGTGCT 552

Qy   1021 TTTGGCATTTGAGCCTGTGATGGGCGGATGAGTTGGTGTGTTTG 1061
                 |||||||||||||||||||||||||||||||||||||||||||
Db   553 CTTCGCGTTGGAACCGGTGATGGCGGCGACGATCGGGNGCCTTCG 593

RESULT 4

LOCUS   DU735587/c
DEFINITION APK13331.g2 HF70_10-07-02 genomic clone HF0070_040C09, genomic survey sequence.
ACCESSION   DU735587
             HF70_10-07-02 uncultured marine microorganism
VERSION   DU735587.1   GI:85745421
KEYWORDS   GSS.
SOURCE   uncultured marine microorganism HF70_10-07-02
  ORGANISM uncultured marine microorganism HF70_10-07-02
            unclassified sequences; environmental samples.
REFERENCE   1 (bases 1 to 916)
  AUTHORS   DeLong,E.F., Preston,C.M., Mincer,T., Rich,V., Hallam,S.J.,
            Frigaard,N.U., Martinez,A., Sullivan,M., Edwards,R., Chisholm,S.W.
            and Karl,D.M.
  TITLE   Comparative genomics reveals ecological trends in stratified
            microbial communities in the ocean's interior
  JOURNAL   Science (2006) In press
COMMENT   Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
            Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson
            and Edward DeLong
            US DOE Joint Genome Institute
            2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
            Tel: 617-253-5271
            Fax: 617-253-2679
            Email: PMRichardson@lbl.gov; delong@mit.edu
            /organism="uncultured marine microorganism HF70_10-07-02"
            /mol_type="genomic DNA"
            /db_xref="taxon:361144"
            /clone="HF0070_040C09"
            /cell_type="marine picoplankton, less than 1.8 um, greater
            than 0.22 um fraction"
            /clone_lib="HF70_10-07-02"
            /note="Vector: pCC1FOS; North Pacific Subtropical Gyre
            Sample Date: 10/7/2002 Coordinates: 22.45 N, 158 W Depth 70 m
            Temperature: 24.93 C Salinity: 35.21 psu Oxygen: 217.4 umol/kg
            Class: fosmid ends.
FEATURES   Location/Qualifiers
     source   1..916

ORIGIN   (Hawaii) picoplankton genomic fosmid DNA library prepared
                from marine picoplankton in the less than 1.6 um, greater
                than 0.22 um fraction. Picoplankton collected at 70 m
                depth on 10/7/2002. Coordinates: 22.45 N, 158 W Sample
                Date: 10/7/2002 Coordinates: 22.45 N, 158 W Depth 70 m
                Temperature: 24.93 C Salinity: 35.21 psu Oxygen: 217.4
                umol/kg"

Qy   Best Local Similarity   54.4%;   Pred. No. 1.7e-11;
     Query Match   5.9%;   Score 80.8; DB 14; Length 916;
     Matches 185; Conservative   0; Mismatches 152; Indels   3; Gaps   1;

Qy   718 ATTGGCATGCGTCTTGGTAGATGAATGGGCGGGGTTGGCGAATTTCAAAAATACGCCAATTTTTG 777
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   372 ATAGTCAGTTTCTTTGTTGACCACTGCGATTGTGTACCGCCGGTACAATGACGCTTTTGT 313

Qy   778 TGCCAAATAATCCACCATGCCATCGTCCGATCAAGATGATCAGCGGAGAGATTAAAAT 837
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   312 TTCGTAGTAGTGCTCCAGGTTGTCGTTAGCCGGATCAAGAGTAGATGATCCGGCCGAAGGTCAAG 253

Qy   838 TGTTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAG 897
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   252 GGTGGCCACATCAGAACGAAAACGACACCTGCACTTTCAAGTTGGTTAACTGGAAAATCTCAAG 193

Qy   898 CACCACCAAAATCCATATTTTGGGATCGTTAATAATTCAAGTGCAGGCGGTTGCCAATATTACC 957
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   192 AATGTAAACCTTTCGCCGATTCATCC---AAAAGATCGAGCGACCAAGGGATACCAAGGTTGCC 136

Qy   958 GCCGACACCAACAATCATGCCTGCAGTGTTTTGCCATCTCGCCTACTAATGTCGTAACAGT 1017
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   135 CCCCACTTTCATGGACAATAGCCAGCATTCCTTCGCGACCCAGTAAATGTCGTTTACTGT 76

Qy   1018 GCTTTTGGCATTTGAGCCTGTGATGGGCGGATGAGTTGGTGTGTTG 1057
                 ||||||||||||||||||||||||||||||||||||
Db   75 ACTTTTGCCATTGGTTCCAGTGATGCGCCATCACCGATGCG 36

RESULT 5

LOCUS   DU773354
DEFINITION APKG3419.g2 HF4000_12-21-03 genomic clone HF4000_[384]008F16, genomic survey
            sequence.
             HF4000_12-21-03 uncultured marine microorganism
ACCESSION   DU773354
VERSION   DU773354.1   GI:85787541
KEYWORDS   GSS.
SOURCE   uncultured marine microorganism HF4000_12-21-03
  ORGANISM uncultured marine microorganism HF4000_12-21-03
            unclassified sequences; environmental samples.
REFERENCE   1 (bases 1 to 1025)
  AUTHORS   DeLong,E.F., Preston,C.M., Mincer,T., Rich,V., Hallam,S.J.,
            Frigaard,N.U., Martinez,A., Sullivan,M., Edwards,R., Chisholm,S.W.
            and Karl,D.M.
  TITLE   Comparative genomics reveals ecological trends in stratified
            microbial communities in the ocean's interior
  JOURNAL   Science (2006) In press
COMMENT   Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
            Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson
            and Edward DeLong
            US DOE Joint Genome Institute
            2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
            Tel: 617-253-5271
            Fax: 617-253-2679
            Email: PMRichardson@lbl.gov; delong@mit.edu
            North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid
            DNA library prepared from marine picoplankton in the less than 1.6
            um, greater than 0.22 um fraction. Sample Date: 12/21/2003
            Coordinates: 22.45 N, 158 W Depth 4000 m Temperature: 1.46 C
            Salinity: 34.69 Oxygen: 147.8 umol/kg
            Class: fosmid ends.
FEATURES   Location/Qualifiers

```
source          1. 1025
                /organism="uncultured marine microorganism"
                HF4000_12-21-03"
                /mol_type="genomic DNA"
                /db_xref="taxon:361148"
                /clone="HF4000 [384]008F16"
                /cell_type="marine picoplankton, less than 1.8 um, greater
                than 0.22 um fraction"
                /clone_lib="HF4000_12-21-03"
                /note="Vector: pCCIFOS; North Pacific Subtropical Gyre
                (Hawaii) picoplankton genomic fosmid DNA library prepared
                from marine picoplankton in the less than 1.6 um, greater
                than 0.22 um fraction. Picoplankton collected at 4000 m
                depth on 10/7/2002. Coordinates: 22.45 N, 158 W. Sample
                Date: 12/21/2003 Coordinates: 22.45 N,158 W
                Temperature: 1.46 C Salinity: 34.69 Oxygen: 147.8 umol/kg"

ORIGIN
Query Match     5.8%;   Score 80.2;  DB 14;   Length 1025;
Best Local Similarity   56.0%;   Pred. No. 2.6e-11;
Matches  173;   Conservative    0;  Mismatches  133;  Indels   3;  Gaps    1;

Qy    762 AAAATACGCAATTTTTTGTGCCAAATAATAATCCACCATGCCATCGTTGCCGATCAAGATGATCAG 821
          |||||| |||||||   |||| |||| | ||||  |||||||||||||||||| |||||||||||
Db     74 AGATGCGTCGCTTGGCCCGCCAGGTAGTCACTGAAGCGGTTGTACGATCCAGATGGTCCT 133

Qy    822 CGGAGAGAGATTTAAAAATTGTTGCCCCCTTGAGCTCGCCAAATTGAGATATGCTCAAGCTGAA 881
          |||||||||||||  |||| || ||| |||| ||  ||||| |||||  ||||||| |||||
Db    134 GGGACACAATTGAGGATCGTGGCCACTCGGCATTCAGGCTGTAGGTGGTTCTCCAACTGGA 193

Qy    882 AACTTGACACAGCTCAAGCACCACCAAAATCATATTTTGGGATCGTTTAATATAATTCAAGTGCAG 941
          |||||  |||||| |||| || ||  |||| | |||  || || ||||   ||| |||||||||
Db    194 AGCTGGACAGCTCCAGGAGCGGTACAGACC--GGCGTCGCTGTGGGCAGCAGATCCAGGCGCCG 250

Qy    942 GCCGTGCCAATATTACCGCCGACACACAACAAATCATGCCTGCATGTTTTGCCATCTCGCCTA 1001
          ||| |||||  ||| ||||| || | | | | |||  |||||| |||| |||||| |||| |
Db    251 GGATTCCCAGGTTGCCGCGCACCCCCGGGTGACGGCGCCAGCCTTCGGCCACTTCACCGA 310

Qy   1002 CTAATGTCGTTAACAGTGCTTTTGGCATTTGAGGCCTGTGTGATGAGTGGGATCGTTGGTGTTTG 1061
          || |||| ||||| | |||| |||| ||||| |||| || ||| |||||| |||| || |||||
Db    311 GCAGGGTGGTGACGGTGCTCTTGGCATTGGAACCGGTGATCGCCACCGCGGGGGCGGCTGG 370

Qy   1062 TTGCTTGAG 1070
          |||||||||
Db    371 CGGCACGGG 379
```

RESULT 6
LOCUS   AG943479/c
DEFINITION Drosophila ananassae DNA, clone: DNB1-017M23.R.fa, genomic survey
            sequence.
ACCESSION   AG943479
VERSION   AG943479.1   GI:58464900
KEYWORDS   GSS.
SOURCE   Drosophila ananassae
  ORGANISM Drosophila ananassae
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1
  AUTHORS   Hattori,M., Toyoda,A., Murakami,K., Kuroki,Y., Fujiyama,A.,
            Toshio,T.K. and Sakaki,Y.
  TITLE   BAC end sequences of library DNB1
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 555)
  AUTHORS   Hattori,M.
  TITLE   Direct Submission
  JOURNAL   Submitted (19-JAN-2005) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)

LOCUS   AG943479   555 bp   DNA   linear   GSS 01-FEB-2005

FEATURES   Location/Qualifiers
     source   1..555
                     /organism="Drosophila ananassae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7217"
                     /clone="DNB1-017M23.R.fa"
                     /clone_lib="DNB1 Drosophila BAC library"

ORIGIN
Query Match   5.6%;   Score 77.2; DB 14;   Length 555;
Best Local Similarity   53.3%;   Pred. No. 1.6e-10;
Matches 163;   Conservative   0; Mismatches 143; Indels   0; Gaps   0;

```
Qy    758 TCAAAAATACGCAATTTTTGTGCCAAATAATCCACCATGCCATCGTTGCCGATCAAGATGA 817
          ||||||||||||||||| ||  |||| ||| |||||||||||||||||| ||||||||||
Db    426 TCAAAAAATATGATGTTTTGCTGTGATAATCCTTGCATATCTCAATGAGCGCTTAAAATGA 367

Qy    818 TCAGCGGAGAGATTTAAAAATTGTTGCCCCCTTGAGCTCGCCAAATTGAGATATGCTCAAGC 877
          ||||  ||| ||||||| ||||   || ||| ||| |  || ||||  ||||| |||||||
Db    366 TCTGGAGTTATATTTAAGTAAACATGAAAATATCAGCTGTAAAATTCTTACGAAACGTTCTAAC 307

Qy    878 TGAAAACTTGACAGCTCAAGCACACCACCAAAATCATATTTTTGGGATCGTTAATAATTCAAGT 937
          |||||| ||||||  ||||| |||  || |||||| |||  || || ||| | ||| |||||
Db    306 ATATAAGAGGACACATCTCTAGAACATATACGCCCTTGATCATTCCAATAAAAGGAAGAGCCAGT 247

Qy    938 GCAGGGCGTGCCAATATTACCGCCGACACACAACAAATCATGCCTGCATGTTTTGCCATCTCG 997
          ||||||   || |||| ||| | |||| | |  ||||  ||| | ||| ||| |||| ||
Db    246 GCAGCGAGGACCTAAATTCCCTCCTCCCTGCCATGCCAGGGAATACACCAGCTTCATTTAAAGATGT 187

Qy    998 CCTACTAATGTCGTTAACAGTGCTTTTGGCATTTGAGGCCTGTGATGGGCGATGGATGGCTGTG 1057
          ||| | ||| |||||| |  | ||||| |||| ||| | || || |||| | ||| || |
Db    186 GTTAATAAACGGTTGTTGTGGATTTCCCATTTGTGCCTGTGCCAACAAAAAACGTGCT 127

Qy   1058 TTTGTT 1063
          ||||||
Db    126 TTTGAT 121
```

RESULT 7
LOCUS   DU748329
DEFINITION ASNC465.b2 HF10_10-07-02 uncultured marine microorganism
            HF10_10-07-02 genomic clone HF0010_001A11, genomic survey sequence.
ACCESSION   DU748329
VERSION   DU748329.1   GI:85758165
KEYWORDS   GSS.
SOURCE   uncultured marine microorganism HF10_10-07-02
  ORGANISM uncultured marine microorganism HF10_10-07-02
            unclassified sequences; environmental samples.
REFERENCE   1 (bases 1 to 977)
  AUTHORS   DeLong,E.F., Preston,C.M., Mincer,T., Rich,V., Hallam,S.J.,
            Frigaard,N.U., Martinez,A., Sullivan,M., Edwards,R., Chisholm,S.W.
            and Karl,D.M.
  TITLE   Comparative genomics reveals ecological trends in stratified
            microbial communities in the ocean's interior

LOCUS   DU748329   977 bp   DNA   linear   GSS 27-JAN-2006

JOURNAL Science (2006) In press
COMMENT Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardeon
and Edward DeLong
US DOE Joint Genome Institute
US DOE Joint Genome Institute
2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
Tel: 617-253-5271
Fax: 617-253-2679
Email: PMRichardson@lbl.gov; delong@mit.edu

FEATURES
source   1..977
/organism="uncultured marine microorganism HF10_10-07-02"
/mol_type="genomic DNA"
/db_xref="taxon:361145"
/clone="HF0010_001A11"
/cell_type="marine picoplankton, less than 1.8 um, greater
than 0.22 um fraction"
/clone_lib="HF10_10-07-02"
/note="vector: pCC1FOS; North Pacific Subtropical Gyre
(Hawaii) picoplankton genomic fosmid DNA library prepared
from marine picoplankton in the less than 1.6 um, greater
than 0.22 um fraction. Picoplankton collected at 10 m
depth on 10/7/2002. Coordinates: 22.45 N, 158 W. Sample
Date: 10/7/2002, Coordinates: 22.45 N, 158 W Depth: 10 m
Temperature: 26.4 C Salinity 35.08 psu Oxygen: 204.6
umol/kg"

ORIGIN

Query Match
Best Local Similarity   49.5%;
Matches 271; Conservative   0; Mismatches 256; Indels 21; Gaps 2;

Score 76.4; DB 14; Length 977;
Pred. No. 3.2e-10;

Qy 338 CCATCAATTGCTCAAATGCTTGAAGTGGTTGAGCCGGATCGTTTGTACCTTTTGAATCATTAAAAATATGCC 397
        ||| || |||| || ||| || ||||| |||| |||| |||||| ||| ||||
Db 346 CTACGAACAGCTGATAACGTTGCCCCCTCATTAGTTGCTTTGGAATCGTTGTAATAGGTT 405

Qy 398 TTGCTTATCAATATATCATCAATATATTCACACGGATCGCGGTAGCCCTTTGAAAGTCTTAAGG 457
        |||| || ||| |||| ||||| || |||||||| ||| ||||| || ||||||| ||||
Db 406 ACGCCCGCAATGACTGCCCACTCTTTCACAGCGGATGAGGCAATCCTTCGAACCCTTTAATA 465

Qy 458 GTATCAAGCACATGGCATCCATGGGAAGAGATTGGCAGCTGTGCCTAAGCAACGTGCAGATAAG 517
        || |||| ||| ||||||||| ||| | ||||||| || |||| ||| ||| ||||
Db 466 GCTCGATACCATTGCATCCACTGGGGATACTCGCCATATCACCAAATCGCTATCGCTCCCACA 525

Qy 518 GCATTGAGTAGGTTATGCCTTGCCTTTGATTTTTAATTGATGGATTGGTTAAAAGTTTGTTT 577
        |||||||||| ||| ||||| ||||| ||||| ||| |||||| ||| || ||| ||||
Db 526 AAAATTCATCAAATGATGCTCACCTTGCAGTGGTATTTCTTCAACCTGCCAACAGATAAATCT 585

Qy 578 TTTTGATGATATAATGCCATACCATCACATCAAAATAAAAAAATCTGCCGTTTGG 637
        ||||| ||||| || ||| ||||||||||| | ||||| |||| ||||||| ||
Db 586 AGACCTCGACTGATAAAATTGCTTCCCTCAGTCTCTACCACACGGTAATGACCTGCTTTA 645

Qy 638 TGGCTAAGTAAGCTATTTGTTGTGTGATAATGGTGTGATTGGGCAACGCTTGTCAGTGTG 697
        ||| |||| |||| ||| | ||||| ||| |||||| ||| ||||||| |||| |||||
Db 646 GGTTCGATCAATACCAAAAACTTTTTCAGACTAG 687

Qy 698 GTCAAGCATTGCTTGGCTTAAATTGGCATCGTTTGGTAGATGAGTGGCGGTTTGGCAAATTT 757
        || |||||||| |||||| |||||| ||| |||| |||||||||| ||||||||||| ||
Db 688 GTGTCCAAGGGCTTAGTTA---GGGATCAATGGGCCGATTAAAAATGATAGAATTCTCGCGCT 744

Qy 758 TCAAAAATACGCAATTTTTGTGTCCAAATAATCACGTGCCGATCAAGAATGA 817
        |||||| | || |||| || |||| |||||| |||||| |||||| ||||
Db 745 TGGTRAAATCCTCTGCGTTATTAAATTGTTGCCGCTGATAATGTCCATGCTGATGAGATGCTCAAGC 804

Qy 818 TCAGCGGAGAGAGATTTAAAATTGTTGCCCCTTGAGCTGCCAAATTAGAAGCATACTCTAAC 877
        ||| |||||||||| || |||||||| || |||| |||| |||||| |||||| ||||
Db 818 AGGAAATCATTAATT... (partial)

RESULT 8
LOCUS   BZ568959   1445 bp   DNA   linear   GSS 17-DEC-2002
DEFINITION pacs2-164_8171.y2 pacs2-164 Pseudomonas aeruginosa genomic clone
pacs2-164_8171, genomic survey sequence.
ACCESSION BZ568959
VERSION BZ568959.1   GI:27202799
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
REFERENCE 1 (bases 1 to 1445)
AUTHORS Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
Burns,J.L., Kaul,R. and Olsen,M.V.
TITLE Whole-Genome Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
JOURNAL J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
Univerity of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: craymond@u.washington.edu
Class: shotgun.

FEATURES
source   1..1445
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pacs2-164_8171"
/clone_lib="pacs2-164"
/note="clinical isolate 2-164 whole genomic shotgun
library."

ORIGIN

Query Match
Best Local Similarity   55.0%;
Matches 137; Conservative   0; Mismatches 112; Indels 0; Gaps 0;

Score 69.8; DB 12; Length 1445;
Pred. No. 2.9e-08;

Qy 327 CCGTGCCAAGCCCCATCAATTGCTGCAATGGTTGAGCCGGATGTTGTACCTTTTGAATCAT 386
        |||||||| || || |||| ||||| |||| |||| || ||| || |||| |||| ||||
Db 87 CGGCACCCAGCCCCTCGATCGCCGCCAGGCGCGCGATCGGGTGTTGGGTGCCCTTTGGAATCGT 146

Qy 387 TAAAAATATGCCTTGCTATCAATATCATCAATATATTCACACGGATCGCGGTAGCCCCTTTGA 446
        ||||| ||||| ||| ||||| |||| ||||| || |||||| ||| |||| || |||||
Db 147 CGTAGTAGCTCACGCCCGTCCGGTACCCACTGGCAGGATGAGCCAGCGGCCCGAAA 206

Qy 447 AAGTTCTTAAGGGTATCAAGACCATGGCATCCATGGGAAGAATTGGCAGCTGTGCCTAAGGCAA 506
        ||||| ||| ||||| || |||| ||| || |||||| |||| | ||||||| || |||||||
Db 207 ACGCCTTCAGCCGCCGCAGCATGGGACCGCGACCGGCAGGCGGCCAGCGCCCA 266

Qy 507 GTGCAGATAAGGCATTGAGTAGGTTATGCCTTGCCTTTGATTTTTAATTGATGGATTGGTA 566
        ||||| |||||| ||||||||| ||| ||||| |||| ||||| |||| |||||| ||||
Db 267 GCGCGCGACGCGTTGGAATAGTTGTGGGCGGCGCCAACGGCA 326

Qy 567 AAAGTTTGT 575
        || |||||
Db 327 GCAGCTTGT 335

RESULT 9
LOCUS   AG952088   AG952088   623 bp   DNA   linear   GSS 01-FEB-2005

Db 805 TCATTTGAAAATGTTTAGAATAGCGGCAACTTTTGCGATCCAAATTAGAAGCATACTCTAAC 864
Qy 878 TGAAAACT 885
        || ||| |
Db 865 TGGAAGCT 872

```
DEFINITION  Drosophila sechellia DNA, clone: DSE1-004A05.F.fa, genomic survey
            sequence.
ACCESSION   AG952088
VERSION     AG952088.1  GI:58473786
KEYWORDS    GSS.
SOURCE      Drosophila sechellia
  ORGANISM  Drosophila sechellia
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 623)
  AUTHORS   Hattori,M., Toyoda,A., Murakami,K., Kuroki,Y., Fujiyama,A.,
            Toshio,T.K. and Sakaki,Y.
  TITLE     BAC end sequences of Library DNB1
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 623)
  AUTHORS   Hattori,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JAN-2005) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa, 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp, URL:http://hgp.gsc.riken.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT     Clones are derived from the BAC library DSE1
            For BAC library availability, please contact Masa-Toshi Yamamoto
            (yamamoto@kit.jp).
            Submitted (30-11-2004) by Masahira Hattori,
            RIKEN, Genomic Sciences Center (GSC);
            1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gsc.riken.jp,Tel:81-45-503-9111,
            Fax:81-45-503-9170)
            This work was done in collaboration with Yamamoto, M-T. Drosophila
            Genetic Resource Center
            Saga Ippongi-cho, Ukyo-ku, Kyoto 616-8354, Japan
            Tel: 81-75-873-2660 FAX: 81-75-861-0881
PRIMERS
            Sequencing : F
LIBRARY
            Vector      : pKS150
            R.Site 1    : SacI.
            R.Site 2    : SacI.
FEATURES             Location/Qualifiers
     source          1..623
                     /organism="Drosophila sechellia"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7238"
                     /clone="DSE1-004A05.F.fa"
                     /clone_lib="DSE1 Drosophila BAC library"
ORIGIN
```

```
Query Match         4.5%;   Score 62.2;  DB 14;  Length 623;
Best Local Similarity  51.0%;   Pred. No. 3.7e-06;
Matches  176;  Conservative  0;  Mismatches 163;  Indels  6;  Gaps  1;

Qy      703 GCATTGCTTGGCTAAAATTGGGCATCGGTCTTGGTAGATGATGGCGGGTTTGGCAATTTTCAAA 762
                ||| | || || |||| ||  ||| | ||||||  |||||||||| | |||||  |||||  |
Db      285 GAATATATCAGCAGTAATCTCATTGTCACGGCACCTATTACGGCACCTCACTCACCGTTTAT 344

Qy      763 AATAGCGCAATTTTTGTGCCAAATAATCCACCATGCCATCGTGCCGATCAGATGATCAGC 822
                ||||||  |||| ||||| ||||| || |  || | |||||| ||| ||||| | |||
Db      285 GAATATATCAGCAGTAATCTCATTGTCACGGCACCTATTACGGCACCTCACTCACCGTTTAT 344

Qy      763 AATAGCGCAATTTTTGTGCCAAATAATCCACCATGCCATCGTGCCGATCAGATGATCAGC 822
                || |||| ||||| |||| ||||| ||  | || | ||||||  || ||||| | |||
Db      345 CAGTTTTAATTTAGTTGCTATGTAGTTCTCCATAACTTCCATGCTCATGTGTGATCTGG 404

Qy      823 GGAGAGATTTAAAAATTGTTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAA 882
                |||||||    ||| ||| | | |   || ||| | ||| | || | | | ||  |||||
Db      405 TGTAATATTGAGCAGTGGCTGAAATGTCCACGTTAATTCATTCATCAACTCCAATTGAAA 464

Qy      883 ACTTGACAGTCAAGCACCACCAAAATCCATATTTTGGATCGTTAATATAATTCAAGTGCAGG 942
                |  |||| | || ||| | ||| |||  || || |||| ||| || || ||||||| |||
Db      465 AGAGGAGAGTTCGATTACATAAATTTCTGCATCTC------TTTCTAGATCCAAAACAGG 518

Qy      943 CGTGCCAATATTACCCCGACACCAAACATCATGCCTGCGCATGTTTTTGCCATCTCGCCTAC 1002
                ||| |||||||  | ||| ||   ||| ||| |||    || | | ||| |||| ||||
Db      519 AACGCCTAAATTCCACCAAATAGCTACTTTTTTCCCTGCAGAGTTTTAATATAGTGCCCTAT 578
```

```
RESULT 10
Qy     1003 TAATGTCGTAACAGTGCTTTTGCGATTTTGAGCCTGTGATGCGAT 1047
                |||| | ||  |||  ||  |  | || |  || |||||||||| |||||||
Db      579 TAGTGATGTAGTGGTTGATTTACCATTTGTTCCTGTGACGCCTAT 623
```

```
DU769542                   928 bp    DNA     linear   GSS 27-JAN-2006
LOCUS       DU769542                   928 bp    DNA     linear   GSS 27-JAN-2006
DEFINITION  HF4000_12-21-03 genomic clone HF4000_[384]003B18, genomic survey
            sequence.
ACCESSION   DU769542
VERSION     DU769542.1  GI:85783725
KEYWORDS    GSS.
SOURCE      uncultured marine microorganism HF4000_12-21-03
  ORGANISM  uncultured marine microorganism HF4000_12-21-03
            unclassified sequences; environmental samples.
REFERENCE   1  (bases 1 to 928)
  AUTHORS   DeLong,E.F., Preston,C.M., Mincer,T., Rich,V., Hallam,S.J.,
            Frigaard,N.U., Martinez,A., Sullivan,M., Edwards,R., Chisholm,S.W.
            and Karl,D.M.
  TITLE     Comparative genomics reveals ecological trends in stratified
            microbial communities in the ocean's interior
  JOURNAL   Science (2006) In press
COMMENT     Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
            Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson
            and Edward DeLong
            Email: PMRichardson@lbl.gov; delong@mit.edu
            US Joint Genome Institute
            North Pacific Subtropical Gyre (Hawaii) picoplankton genomic fosmid
            DNA library prepared from marine picoplankton in the less than 1.6
            um, greater than 0.22 um fraction. Sample Date: 12/21/2003
            Coordinates: 22.45 N, 158 W Depth 4000 m Temperature: 1.46 C
            Salinity: 34.69 Oxygen: 147.8 umol/kg
            Class: fosmid ends.
FEATURES             Location/Qualifiers
     source          1..928
                     /organism="uncultured marine microorganism
                     HF4000_12-21-03"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:361148"
                     /clone="HF4000_[384]003B18"
                     /clone_lib="HF4000_12-21-03"
                     /note="Vector: pCC1FOS; North Pacific Subtropical Gyre
                     (Hawaii) picoplankton genomic fosmid DNA library prepared
                     from marine picoplankton in the less than 1.6 um, greater
                     than 0.22 um fraction. Picoplankton collected at 4000 m
                     depth on 10/7/2002. Coordinates: 22.45 N,158 W. Sample
                     Date: 12/21/2003 Coordinates: 22.45 N, 158 W Depth 4000 m
                     Temperature: 1.46 C Salinity: 34.69 Oxygen: 147.8 umol/kg"
ORIGIN
```

```
Query Match         4.3%;   Score 59.6;  DB 14;  Length 928;
Best Local Similarity  50.3%;   Pred. No. 2.3e-05;
Matches  146;  Conservative  0;  Mismatches 144;  Indels  0;  Gaps  0;

Qy      758 TCAAAAAATACCGCAATTTTTGTGCCAAATAAATCACCATGCCATCGTGCCGATCAAGATGA 817
                || |||    || ||| || | | | ||| ||||  | || || ||  |||| ||  ||
Db      143 TAAAAGATCATCTTCTTGGGCGGTATTTGCCCTTGAATCGTCCAAATGCGGGCGTAACGGTG 202

Qy      818 TCAGCCGGAGAGATTTAAAATGTTGCCCTTGAGCTGCCAAATTTGAGATATGCCCAAGC 877
                || || || | ||||  | | |||| || | || ||| | ||| ||  ||| |||||
Db      203 TCGGGGCGTGATGTTCAACAGCACAGCAGGACGTCGAAGGTGATCGAACGGGTCAAGTTCGGAGC 262

Qy      878 TGAAAACTTGACAGCTCAAGCACCACCAAATCCATATTTTGGATCGTTAATAATTCAAGT 937
                |||   ||  |||  |   ||  | || |  | || | |||  || |  ||  ||||
Db      878 TGAAAACTTGACAGCTCAAGCACCACCAAATCCATATTTTGGATCGTTAATAATTCAAGT 937
```

RESULT 11

```
LOCUS       CL691264/c
DEFINITION  PRI0156a_A05_2 - PRI0156a.BR (628) Mixed stage fosmid library of P.
            pacificus var. California Pristionchus pacificus genomic, genomic
            survey sequence.
ACCESSION   CL691264
VERSION     CL691264.1  GI:50213172
KEYWORDS    GSS.
SOURCE      Pristionchus pacificus
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 628)
  AUTHORS   Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
  TITLE     AppaDB: an AcedB database for the nematode satellite organism
            Pristionchus pacificus
  JOURNAL   Nucleic Acids Res. 32 (1), D421-D422 (2004)
   PUBMED   14681447
COMMENT     Contact: Sommer RJ
            Evolutionary Biology
            Max-Planck-Institute for Developmental Biology
            Spemannstr. 37-39, Tuebingen D-72076, Germany
            Tel: 0049707601498
            Fax: 00497071601371
            Email: ralf.sommer@tuebingen.mpg.de
            This library was generated at Caltech, Pasadena, USA and end
            sequenced at Vancouver, Canada.
            Seq primer: T7
            Class: fosmid ends.
FEATURES             Location/Qualifiers
     source          1..628
                     /organism="Pristionchus pacificus"
                     /mol_type="genomic DNA"
                     /strain="California"
                     /clone_lib="Mixed stage fosmid library of P. pacificus
                     var. California"
                     /note="Vector: pEpifos-5 fosmid vector"
ORIGIN
```

Query Match   4.1%; Score 56.8; DB 13; Length 628;
Best Local Similarity   48.0%; Pred. No. 0.00014;
Matches 246; Conservative   0; Mismatches 242; Indels 24; Gaps 2;

```
Qy     1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATATTTTAAATTGATC  60
           ||||  | | |||| || | |||| ||| || | | ||| | | | || |  |||||
Db   494 TAACTCCTTCGCACGACGGGCAAACTCATTGCCTCGTTGTTCAAAGTTCTTGAACTGATC 435

Qy    61 CATGCTAGCACAACAGGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 120
          || | ||   | ||  ||| || ||  ||  || |  |  | | | |   | | ||  ||
Db   434 AAGGCTGGCACAGGCTGGGGAGCAGAGAACCAT----------ATCGCCCGGCTG     390

Qy   121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCAAAATGTGCATCACCTTT 180
           || ||   ||| | |||  | | | | |  | | ||  ||| | |  | | |||| |||
Db   389 AACACGCGAGCCAGCCACGCATCGCCTGTTCATAGTTCGGTTGTTGTTCTGCCACTTC   330

Qy   181 GAGTAAATCTTGCTCAATTATCGGGGCATCTTGACGGATTAAATACACACTATCCACCCA 240
          || | ||  | ||  || | ||  || | ||||  ||| |||| | ||| | |||| |||
Db   329 CGGGCCGTAGCGCCGCCAGCTGCGCCGGTCACGACGGAAACAATACAGAGGTAGCGTTATC 270

Qy   241 CTGCATAACATATCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCGCCACCAAAAT  300
          |||| | |  | |     | || |  |     || |  |  |||  |  | || |
Db   241 CTGCATAACATATCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCGCCACCAAAAT  334
```

RESULT 12

```
LOCUS       DR032465/c
DEFINITION  bda020270D15.ab1_061 Antrodia cinnamomea cDNA, mRNA sequence.
            cinnamomea cDNA, mRNA sequence.          474 bp    mRNA    linear    EST 31-MAY-2005
ACCESSION   DR032465
VERSION     DR032465.1  GI:66834715
KEYWORDS    EST.
SOURCE      Antrodia cinnamomea
  ORGANISM  Antrodia cinnamomea
            Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
            Aphyllophorales; Antrodia.
REFERENCE   1  (bases 1 to 474)
  AUTHORS   Tzean,S.S., Hwang,S.T., Song,H.C., Liou,R.F., Yehk,H.H. and
            Shiao,M.S.
  TITLE     Analysis of expressed sequence tags from Antrodia cinnamomea, a
            polypore causing brown rot of Cinnamomum kanehirai
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Tzean SS
            Department of Plant Pathology and Microbiology
            National Taiwan University
            No.1, Section 4, Roosevelt Road, Taipei 10617, Taiwan
            Tel: 886 2 3366 4595
            Fax: 886 2 2362 0639
            Email: sst@ntu.edu.tw
            PCR Primers
            FORWARD: 5' GTAAAACGACGGCCAGT 3'.
            BACKWARD: 5' CACACAGGAAACAGCTATGACCAT 3'.
FEATURES             Location/Qualifiers
     source          1..474
                     /organism="Antrodia cinnamomea"
                     /mol_type="mRNA"
                     /strain="SMV 1"
                     /db_xref="taxon:279009"
                     /sex="Heterokaryon"
                     /cell_type="Mycelium"
                     /lab_host="Cinnamomum kanehirai"
                     /clone_lib="Antrodia cinnamomea cDNA Library"
                     /note="Vector: ZAP Express Vector"
ORIGIN
```

Query Match   4.0%; Score 55; DB 9; Length 474;
Best Local Similarity   54.1%; Pred. No. 0.00042;
Matches 112; Conservative   0; Mismatches 95; Indels   0; Gaps   0;

```
Qy  1171 AATGCCGGGGGCTGATGATGATCTGATCGGCTGAGCTTAAGAGTTCACTGTGTCCAAGCTGCC 1230
          ||| | ||  | ||| ||| ||| ||  | || | ||| ||  ||  || | |  |||| ||
Db   453 AATACCGGGACTGGGCGACAATCAGATCTGCCGCATCAGCCATTCATCAGACTGCC   394

Qy  1231 AAAATGACAATGAAGCGCCTGTGGGCAGTTTTATCTGCCCAAGGTAGGATTGGGATTACCATC 1290
          ||| |   ||   | ||  ||  |  || |  | || ||  ||  |   | |||| | |||
Db   393 CGTGTGGCGGTTCTACGGCTTCGGGTGGTTATTATTCAGGCCAGGCGGTGTCTATAGCGGTATC 334
```

```
Qy    1291  GATGATGCTAACTTTGTGACCTGATTGACCAAAAAATTCACACGCTGACAGTCCAGAGAGCT  1350
            ||||||  ||||| | ||  ||| |||||||||||||||||||  || |||||||||||||
Db     333  CATAACGCGCGGGCGTCACACCGGCGAGCGAGAGAAAAGTCCACGCAGGAAAAGCCCGGTGAG  274

Qy    1351  GCCCAAACCGACGACGGGCATATTTTTTT  1377
            ||||||||  ||  ||||||||||||||
Db     273  GCCCAGGCCGGATAATGACGACACATTTTT  247
```

RESULT 13

DU732353/c

<pre>
LOCUS   DU732353   1024 bp   DNA   linear   GSS 27-JAN-2006
DEFINITION HF70_10-07-02 uncultured marine microorganism
            HF70_10-07-02 genomic clone HF0070_025B08, genomic survey sequence.
ACCESSION   DU732353
VERSION   DU732353.1 GI:85742187
KEYWORDS   GSS.
SOURCE   uncultured marine microorganism HF70_10-07-02
  ORGANISM uncultured marine microorganism HF70_10-07-02
            unclassified sequences; environmental samples.
REFERENCE   1 (bases 1 to 1024)
  AUTHORS   DeLong,E.F., Preston,C.M., Mincer,T., Rich,V., Hallam,S.J.,
            Frigard,N.U., Martinez,A., Sullivan,M., Edwards,R., Chisholm,S.W.
            and Karl,D.M.
  TITLE   Comparative genomics reveals ecological trends in stratified
            microbial communities in the ocean's interior
  JOURNAL   Science (2006) In press
COMMENT   Contact: Susan Lucas, Alex Copeland, Sam Pitluck, Alla Lapidus,
            Kerrie Barry, Tijana Glavinadelrio, David Bruce, Paul Richardson
            and Edward DeLong
            US DOE Joint Genome Institute
            2800 Mitchell Drive B100, Walnut Creek, CA 94598-1698, USA
            Tel: 617-253-5271
            Fax: 617-253-2679
            Email: PMRichardson@lbl.gov; delong@mit.edu
            Sample Date: 10/7/2002 Coordinates: 22.45 N, 158 W Depth 70 m
            Temperature: 24.93 C Salinity: 35.21 psu Oxygen: 217.4 umol/kg
            Class: fosmid ends.
FEATURES   Location/Qualifiers
     source   1..1024
                     /organism="uncultured marine microorganism
                     HF70_10-07-02"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:361144"
                     /clone="HF0070_025B08"
                     /cell_type="marine picoplankton, less than 1.8 um, greater
                     than 0.22 um fraction"
                     /clone_lib="HF70_10-07-02"
                     /note="Vector: pCC1FOS; North Pacific Subtropical Gyre
                     (Hawaii) picoplankton genomic fosmid DNA library prepared
                     from marine picoplankton in the less than 1.6 um, greater
                     than 0.22 um fraction. Picoplankton collected at 70 m
                     depth on 10/7/2002. Coordinates: 22.45 N, 158 W Sample
                     Date: 10/7/2002 Coordinates: 22.45 N, 158 W Depth 70 m
                     Temperature: 24.93 C Salinity: 35.21 psu Oxygen: 217.4
                     umol/kg"
ORIGIN
</pre>

```
Query Match       3.9%;   Score 54.2; DB 14; Length 1024;
Best Local Similarity   51.4%;   Pred. No. 0.0088;
Matches  125;   Conservative    0;  Mismatches  118;  Indels    0; Gaps    0;

Qy     333  CAAGCCCATCAATTGCTGCAAATGGTTGAGCCGATGGTTGTACCTTTTGAATCATTAAAAT  392
            ||| |||| |||||  ||||| ||||||| ||||| ||||  ||||||| ||| ||||||
Db     557  CCAGGCCCTCCAGCGGCCCCGTCGCACCGGATATTGGGTCCTTTGCTGTGTCATCGAATGA  498

Qy     393  ATGCCTTGCTATCATCAAATATCACACGGATGCGGCGTAGCCCTTTGAAAGTCT  452
            ||||||||||  |||||||||||||||||||||||||||||||  ||||||||
Db     497  AGGCGCACGCCCTTGACTTGACTGATTAGCTGCGGCCGATGGGGCCAGAACCTCGGCACTGCT  438

Qy     453  TAAGGGGTATCAAAGCAATGGCATCCAATGGGAAGAGATTGGCAGCGTGGGGGCATGGGGCAG  512
            |||||  |||||||||||||||||  |||||  ||||||  ||||| ||||||||||||||
Db     453  AGCGGCACGCCCTTGACTTGACTGATTAGCTGCGGCCGATGGGGCCAGAACCTCGGCACTGCT  438
```

```
Qy     437  TTGCTCCCTCGATCAGTCCCTCGAGGGGCATTTCAACGCGCCGCTTGACACAGCCG  378
            |||| || |||  |||||| || ||||| ||  ||| |||||||||||||||||
Db     437  TTGCTCCCTCGATCAGTCCCTCGAGGGGCATTTCAACGCGCCGCTTGACACAGCCG  378
```

RESULT 14

B255906/c

<pre>
LOCUS   B2559806   1171 bp   DNA   linear   GSS 17-DEC-2002
DEFINITION pacs2-164_1863.x1 pacs2-164 Pseudomonas aeruginosa genomic clone
            pacs2-164_1863, genomic survey sequence.
ACCESSION   B2559806
VERSION   B2559806.1 GI:27176698
KEYWORDS   GSS.
SOURCE   Pseudomonas aeruginosa
  ORGANISM Pseudomonas aeruginosa
            Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
            Pseudomonadaceae; Pseudomonas.
REFERENCE   1 (bases 1 to 1171)
  AUTHORS   Spencer,D.H., Raymond,C.K., Smith,E.E., Sims,E.E., Hastings,M.,
            Burns,J.L., Kaul,R. and Olson,M.V.
  TITLE   Whole-Genome-Sequence variation among multiple isolates of
            Pseudomonas aeruginosa
  JOURNAL   J. Bacteriol. (2002) In press
COMMENT   Contact: Chris K. Raymond
            Genome Center
            University of Washington
            Box 352145, Seattle, WA 98105-2145, USA
            Tel: 2062216954
            Fax: 2068857244
            Email: craymond@u.washington.edu
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..1171
                     /organism="Pseudomonas aeruginosa"
                     /mol_type="genomic DNA"
                     /strain="2-164"
                     /db_xref="taxon:287"
                     /clone="pacs2-164_1863"
                     /clone_lib="pacs2-164"
                     /note="clinical isolate 2-164 Whole genomic shotgun
                     library."
ORIGIN
</pre>

```
Query Match       3.8%;   Score 52.6; DB 11; Length 1171;
Best Local Similarity   49.0%;   Pred. No. 0.0026;
Matches  170;   Conservative    0;  Mismatches  174;  Indels    3; Gaps    1;

Qy     513  ATAAGGCATTGAGTAGGTTATAGGCTTGCCTTTGATTTTAATTGATGGATTGGTAAAGTT  572
            || ||||||||| || |||||| ||||||||||| ||  |||| |||||  ||| |||||
Db     377  CCAGCACATTCAAGTGGTTATGGCCTGCCTTCCAGCTGGACCGCCACAGCAGTC  318
```

```
Qy     573  TGT  575
            |||
Db     317  TCT  315
```

```
Qy     177  CTTTGAGTAAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACTATCCA  236
            |||| || ||  ||||| |||||||| ||| ||| |||| |||| ||||||| |||||||
Db     396  CGTTGCCCAGTGCCTGTCAATCAGTCTAGGTGTCACGCACAAGCAGTACCACCGCCCGGC  337
```

```
Qy     237  CCCACTGCATAACATATGCTGACAGCAGCTCACTAAAAATCTTGACCTTTACCAAGCCACCA  296
            |||||||  ||||||||||||||| |||||||||||  |||||||  |||| ||||||||
Db     336  AGAAGGCCGCGCAACCGGCGTCGCCCAGTTCATGAAAAATCGGGCCCCTTGCCGGTCTCCGCGG  277
```

```
Qy     297  AAATCAATGCCAGTGATGATATGCCATAAAACCGTGCCAAGCCCCATCAATTGCTGCAATGG  356
            |||||| |  |  | || ||||||| ||||||||||||||| |||||| ||||  ||||||
Db     276  CGATCAGCACCAG---CTTGCGCGTCGATGTCGATGTCGGGCACCACCCCTCGATCGGCCAGGG  220
```

```
Qy     357  TTGAGCCGATGGTTGTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCATCATCAA  416
            |||||| ||| ||  ||||||| |||||||||||  ||||||||||||||||  |||||
Db     219  CGGCGCGCACGTTGGTGGGTCGGGTGGCCTTGGAATCGGTCGTTAGTAGCGTCACGCCCGGTA  160
```

```
Qy     417  TATATTCACAGCGGATGCGGGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAAGCATGGCATCCA  476
            ||||||||||| ||||| |||||||||||||||||||| ||||| |||||||||  |||||
Db     159  CCCACTGGCAGAATGAGCCAGTCCGCAAGACGCTTTCAGCGCGCCGCAGCATGGCCGTAAA  100
```

```
Qy     477  TCGGAAGAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTG  523
            ||||  |||||||| |||| || |||||||||| | ||| ||||| |||
Db     477  TCGGAAGAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTG  523
```

```
Db         99  ACGGCAGGGGACCGACTGGCCCAGTGCCAGTCGCTCGGCGAGCGCGTTG 53
               |||.|  ||||||   |||||  |||||||||||  |  ||| |||| |||

RESULT 15
LOCUS       CZ545153                 838 bp    DNA     linear   GSS 13-MAY-2005
DEFINITION  SRAA-aad60c01.g1 Strongyloides ratti whole genome shotgun library
            (SRAAGSS 004) Strongyloides ratti genomic, genomic survey sequence.
ACCESSION   CZ545153
VERSION     CZ545153.1  GI:64677289
KEYWORDS    GSS.
SOURCE      Strongyloides ratti
  ORGANISM  Strongyloides ratti
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Panagrolaimoidea; Strongyloididae; Strongyloides.
REFERENCE   1  (bases 1 to 838)
  AUTHORS   Mitreva,M., McCarter,J.P., Thompson,F., Viney,M., Pape,D.,
            Ritter,E., Martin,J., Wylie,T., Dante,M., Waterston,R.H.,
            Clifton,S.W. and Wilson,R.
  TITLE     Genome Survey sequences from the rat parasitic nematode
            Strongyloides ratti
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Mitreva M
            Washington University in St. Louis
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: nematode@watson.wustl.edu
            Genomic DNA was provided by Fiona Thompson
            (F.Thompson@bristol.ac.uk) and Mark Viney
            (Mark.Viney@bristol.ac.uk) at the University of Bristol, Bristol,
            UK.
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1..838
                     /organism="Strongyloides ratti"
                     /mol_type="genomic DNA"
                     /strain="Isofemale line ED321 heterogonic"
                     /db_xref="taxon:34506"
                     /dev_stage="infective larval stage (iL3)"
                     /lab_host="GS10"
                     /clone_lib="Strongyloides ratti whole genome shotgun
                     library (SRAAGSS 004)"
                     /note="Vector: pOTW13; Site_1: BstXI; Site_2: BstXI;
                     Strongyloides ratti genomic DNA was randomly sheared,
                     end-repaired and size fractioned to enrich for 2-4 kb
                     fragments. Genomic DNA was provided by Fiona Thompson
                     (F.Thompson@bristol.ac.uk) and Mark Viney
                     (Mark.Viney@bristol.ac.uk) at the University of Bristol,
                     Bristol, UK. Sequencing by Washington University Genome
                     Sequencing Center, St. Louis, MO."
ORIGIN
Query Match       3.6%;  Score 50.2;  DB 13;  Length 838;
Best Local Similarity   62.2%;  Pred. No. 0.012;
Matches  79;  Conservative  0;  Mismatches  48;  Indels  0;  Gaps  0;

Qy      946 GCCAATAATTACCCGCGACACCAACAAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAA 1005
                |||||||||||||  |||||||  |  ||||   |||||||||||||||  ||||||||||
Db       76 GCCAATATTTCCGCCTAAACACATTCAAATCCTTCATTTTTCAGGATGTGATAGATTAA 135

Qy     1006 TGTCGTAACAGATGCTTTGGCATTTGAGCCTGTGATGGCGATGATGGTGTGTTTGTTTGC 1065
                ||||||||||||| ||||| |||||||||||||| || |||||||| ||||  |||| ||
Db      136 AGACGTTGTCGTTCGTTTTTTCCGTTGCTTCCTGTGATCGCAATGATTTTCGCATCGGTAAA 195

Qy     1066 TTGAGTT 1072
                ||| |||
Db      196 TTCAGAT 202
```

Search completed: May 18, 2006, 11:52:44

Job time : 6248 secs

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

```
                    GenCore version 5.1.8
                    Copyright (c) 1993 - 2006  Bioacceleration Ltd.

Run on:             May 18, 2006, 09:25:31 ; Search time 255 Seconds
                                             (without alignments)
                                             10126.001 Million cell updates/sec

Title:              US-10-672-787-35_COPY_11357_12736
Perfect score:      1380
Sequence:           1 taaagcgtgaaccatgttaa.............acgacggcatattttttga 1380

Scoring table:      IDENTITY_NUC
                    Gapop 10.0 , Gapext 1.0

Searched:           1403666 seqs, 935554401 residues

Total number of hits satisfying chosen parameters:    2268460

Minimum DB seq length: 22
Maximum DB seq length: 2000000000

Post-processing:    Minimum Match 0%
                    Maximum Match 100%
                    Listing first 45 summaries

Database :          Issued Patents_NA:*
                    1:  /EMC_Celerra_SID53/ptcdata/2/ina/1_COMB.seq:*
                    2:  /EMC_Celerra_SID53/ptcdata/2/ina/5_COMB.seq:*
                    3:  /EMC_Celerra_SID53/ptcdata/2/ina/6A_COMB.seq:*
                    4:  /EMC_Celerra_SID53/ptcdata/2/ina/6B_COMB.seq:*
                    5:  /EMC_Celerra_SID53/ptcdata/2/ina/7_COMB.seq:*
                    6:  /EMC_Celerra_SID53/ptcdata/2/ina/H_COMB.seq:*
                    7:  /EMC_Celerra_SID53/ptcdata/2/ina/PCTUS_COMB.seq:*
                    8:  /EMC_Celerra_SID53/ptcdata/2/ina/PP_COMB.seq:*
                    9:  /EMC_Celerra_SID53/ptcdata/2/ina/RE_COMB.seq:*
                    10: /EMC_Celerra_SID53/ptcdata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1380	100.0	96109	3	US-09-596-002-35	Sequence 35, Appl
2	1370.4	99.3	1410	3	US-09-544-236-673	Sequence 673, App
3	226.2	16.4	1416	3	US-09-328-352-1282	Sequence 1282, App
4	160.4	11.6	1830121	3	US-09-557-884-1	Sequence 1, Appli
5	160.4	11.6	1830121	3	US-09-643-990A-1	Sequence 1, Appli
6	160.4	11.6	1830121	3	US-10-158-865-1	Sequence 1, Appli
7	125.2	9.1	1323	3	US-09-543-681A-2830	Sequence 2830, Ap
8	117	8.5	1371	3	US-09-252-991A-7702	Sequence 7702, Ap
9	117	8.5	1401	3	US-09-252-991A-7861	Sequence 7861, Ap
10	117	8.5	1450	3	US-09-701-229-1	Sequence 1, Appli
11	100.6	7.3	1416	3	US-09-489-0399A-5247	Sequence 5247, Ap
12	84.8	6.1	640681	3	US-09-790-988-1	Sequence 1, Appli
13	83	6.0	1368	3	US-09-107-532A-1574	Sequence 1574, Ap
14	69.8	5.1	567	3	US-09-252-991A-7787	Sequence 7787, Ap
15	67	4.9	2402	3	US-09-221-017B-881	Sequence 881, App
16	61.4	4.4	1372	3	US-09-530-836-1	Sequence 1, Appli
17	58.6	4.2	1242	3	US-09-147-928-1	Sequence 1, Appli
18	53.6	3.9	2423	3	US-08-714-918-86	Sequence 86, Appl
19	53.6	3.9	2423	3	US-09-265-315-86	Sequence 86, Appl
20	53.6	3.9	2423	3	US-09-265-315-86	Sequence 86, Appl
21	53.6	3.9	2423	3	US-09-266-417-86	Sequence 86, Appl
22	53.6	3.9	2423	3	US-09-528-709-86	Sequence 86, Appl
23	53.6	3.9	2423	3	US-09-527-745-86	Sequence 86, Appl
24	51.8	3.8	1352	5	US-09-974-300-1685	Sequence 1685, Ap
25	49.4	3.6	536	3	US-09-147-928-3	Sequence 3, Appli
26	49.4	3.6	699	3	US-09-134-000C-741	Sequence 741, App
27	48.6	3.5	1320	3	US-08-781-986A-80	Sequence 80, Appl
28	48.6	3.5	1188	3	US-09-710-279-655	Sequence 655, App
29	48.2	3.5	3334	3	US-09-710-279-3574	Sequence 3574, Ap
30	48.2	3.4	1374	3	US-09-134-001C-1001	Sequence 1001, App
31	47	3.4	405	3	US-09-974-300-6173	Sequence 6173, Ap
32	45.6	3.3	717	3	US-09-974-300-997	Sequence 997, App
33	45.6	3.3	1368	3	US-09-936-165A-78	Sequence 78, Appl
34	42.6	3.1	10486	3	US-09-902-540-9337	Sequence 9337, App
35	42.6	3.1	1407	3	US-09-902-540-997	Sequence 997, App
36	42	3.0	6693	3	US-08-961-527-195	Sequence 195, App
37	41.8	3.0	4010	3	US-09-710-279-3540	Sequence 3540, Ap
38	41	3.0	1141	3	US-10-113-113-3	Sequence 3, Appli
39	41	3.0	1275	3	US-09-806-708B-22	Sequence 22, Appl
40	40.8	3.0	1350	2	US-08-665-435A-1	Sequence 1, Appli
41	40.8	2.9	1350	2	US-08-665-435A-3	Sequence 3, Appli
42	40.2	2.9	1350	2	US-08-843-309-3	Sequence 3, Appli
43	40.2	2.9	1350	2	US-08-843-309-1	Sequence 1, Appli
44	40.2	2.9	1353	2	US-09-583-110-197	Sequence 197, App
45	40.2	2.9	1353	3	US-09-583-110-197	Sequence 197, App

ALIGNMENTS

RESULT 1
US-09-596-002-35

```
; Sequence 35, Application US/09596002
; Patent No. 6632636
; GENERAL INFORMATION:
; APPLICANT:     Lagace, Robert, E.
; APPLICANT:     Patterson, Chandra
; APPLICANT:     Berg, Kim, L.
; TITLE OF INVENTION:  NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
; FILE REFERENCE:  PM-0008-4 US
; CURRENT APPLICATION NUMBER:  US/09/596,002
; CURRENT FILING DATE:  2000-06-16
; PRIOR APPLICATION NUMBER:  60/140,121
; PRIOR FILING DATE:  1999-06-18
; NUMBER OF SEQ ID NOS:  41
; SOFTWARE:  PERL Program
; SEQ ID NO 35
;   LENGTH: 96109
;   TYPE: DNA
;   ORGANISM: M. catarrhalis
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte template ID No. 6632636_35
; PUBLICATION INFORMATION:
US-09-596-002-35

Query Match             100.0%;  Score 1380;  DB 3;  Length 96109;
Best Local Similarity   100.0%;
Matches 1380;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy          1 TAAAGCGTGAACCATGTTAACAACATTATCACCTCGCTCATTATAATTTTTAAAATTGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11357 TAAAGCGTGAACCATGTTAACAACATTATCACCTCGCTCATTATAATTTTTAAAATTGATC 11416

Qy         61 CATGGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGGTGCTTTGACTGGGCTGCCAA 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11417 CATGGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGGTGCTTTGACTGGGCTGCCAA 11476

Qy        121 TACTCGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11477 TACTCGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 11536

Qy        181 GAGTAAATCTTTGCTCAATTATCGGGGGCATCTTGACCGGATTAAAATACACACTATCCAAACCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      11537 GAGTAAATCTTTGCTCAATTATCGGGGGCATCTTGACCGGATTAAAATACACACTATCCACCCA 11596
```

```

```
Qy   241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11597 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 11656

Qy   301 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11657 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 11716

Qy   361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11717 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 11776

Qy   421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11777 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 11836

Qy   481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11837 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 11896

Qy   541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCCATACC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11897 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCCATACC 11956

Qy   601 ATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTTGTTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11957 ATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTTGTTGT 12016

Qy   661 GATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12017 GATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 12076

Qy   721 GGCATCGTCTTGGTAGATGATGGCGGTTTGGCAATTTTCAAAAATACGCAATTTTTGTGC 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12077 GGCATCGTCTTGGTAGATGATGGCGGTTTGGCAATTTTCAAAAATACGCAATTTTTGTGC 12136

Qy   781 CAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTGT 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12137 CAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTGT 12196

Qy   841 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCAC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12197 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCAC 12256

Qy   901 CACCAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTACCGCC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12257 CACCAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTACCGCC 12316

Qy   961 GACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGCT 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12317 GACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGCT 12376

Qy   1021 TTTGGCATTTGAGCCTGTGATGGCGATGATTGGTGTGTTTGTTGCTTGAGTTAGATCTCG 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12377 TTTGGCATTTGAGCCTGTGATGGCGATGATTGGTGTGTTTGTTGCTTGAGTTAGATCTCG 12436

Qy   1081 TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTTT 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12437 TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTTT 12496

Qy   1141 GGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGGC 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12497 GGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGGC 12556

Qy   1201 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTTT 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12557 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTTT 12616

Qy   1261 ATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGAC 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12617 ATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGAC 12676

Qy   1321 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCATATTTTTTTGA 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   12677 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCATATTTTTTTGA 12736

RESULT 2
US-09-540-236-673/c
; Sequence 673, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 673
;   LENGTH: 1410
;   TYPE: DNA
;   ORGANISM: M.catarrhalis
US-09-540-236-673

  Query Match   99.3%; Score 1370.4; DB 3; Length 1410;
  Best Local Similarity   99.6%; Pred. No. 0;
  Matches 1374; Conservative   0; Mismatches   6; Indels   0; Gaps   0;

Qy   1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1392 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 1333

Qy   61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 120
              |||||||||||||||||| |||||||||||||||||||||||||||||||||||||||||
Db   1332 CATGCTAGCACAAGCAGGCGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 1273

Qy   121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1272 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 1213

Qy   181 GAGTAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACTATCCACCCA 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1212 GAGTAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACTATCCACCCA 1153

Qy   241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1152 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 1093

Qy   301 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1092 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 1033

Qy   361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db   1032 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCCATCAATATCATCAATATA 973

Qy   421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   972 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 913

Qy   481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   912 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 853

Qy   541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCCATACC 600
              |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db   852 TTTGATTTTTAATTGATGGATTGGTAAAAGCTTGTTTTTTTGATGATATAATGCCATACC 793

Qy   601 ATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTTGTTGT 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   792 ATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTTGTTGT 733

Qy   661 GATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   732 GATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAATT 673
```

```
Qy          721 GGCATCGTCTTGGTAGATGATGGCGGTTTGGCAATTTTCAAAAATACGCAATTTTTGTGC 780
                ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db          672 GGCATCGTCTTGGCAGATGATGGCGGTTTGGCAATTTTCAAAAATACGCAATTTTTGTGC 613

Qy          781 CAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTGT 840
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          612 CAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTGT 553

Qy          841 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCAC 900
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          552 TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCAC 493

Qy          901 CACCAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTACCGCC 960
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          492 CACCAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTACCGCC 433

Qy          961 GACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGCT 1020
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          432 GACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGCT 373

Qy         1021 TTTGGCATTTGAGCCTGTGATGGCGATGATTGGTGTGTTTGTTGCTTGAGTTAGATCTCG 1080
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          372 TTTGGCATTTGAGCCTGTGATGGCGATGATTGGTGTGTTTGTTGCTTGAGTTAGATCTCG 313

Qy         1081 TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTTT 1140
                ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||| ||
Db          312 TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGGATGCCTTGAGCCTT 253

Qy         1141 GGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGGC 1200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          252 GGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGGC 193

Qy         1201 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTTT 1260
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          192 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTTT 133

Qy         1261 ATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGAC 1320
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          132 ATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGAC 73

Qy         1321 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCATATTTTTTTGA 1380
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           72 CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCATATTTTTTTGA 13

RESULT 3
US-09-328-352-1282/c
; Sequence 1282, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
;  APPLICANT: Gary L. Breton et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
;  TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: GTC99-03PA
;  CURRENT APPLICATION NUMBER: US/09/328,352
;  CURRENT FILING DATE: 1999-06-04
;  NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1282
;   LENGTH: 1416
;   TYPE: DNA
;   ORGANISM: Acinetobacter baumannii
US-09-328-352-1282

  Query Match             16.4%;  Score 226.2;  DB 3;  Length 1416;
  Best Local Similarity   52.1%;  Pred. No. 1.2e-58;
  Matches  718;  Conservative    0;  Mismatches  593;  Indels   66;  Gaps    7;

Qy            2 AAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATCC 61
                || |  |  ||   |  ||||||||  | ||| || ||||||||| |||||||   |||
Db         1409 AACGAATTGACGCAGGCAACAAACTGCTGACCACGGTCATTATAACTTTTAAACATATCA 1350

Qy           62 ATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAAT 121
                |  || ||||| || |||||         |||  || |    |||   || |      |
Db         1349 AAACTTGCACATGCTGGTGAT--------AGCAATACCACATCTTCAGCTTGTGTTTCAC 1298

Qy          122 ACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTTG 181
                  || |  | |  |    |     | |    |  |  |||  |||| |   |  |   |
Db         1297 GTTGACACAGCTCGACAGCTTCTTTAAGCGTTGCTGCATGTAAAATTTTAGTTGCGCCT- 1239

Qy          182 AGTAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACTATCCACCCAC 241
                 | |   ||||| || || |  || ||||||| ||| || ||   ||||   |   |  |
Db         1238 TGAATGGCTTGTTCGATGACAGGCGCATCTTCACCAATCAATACCACAACTTTGGCATAT 1179

Qy          242 TGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAATC 301
                |     | | | |     |    |  |||||||| ||||| ||  | ||||| |||||
Db         1178 TTTTCGATAGAAGAACGTAAAGGAGAAAAATCTTGTCCTTTGCCCTGTCCACCTAAAATA 1119

Qy          302 AATGCCA------GTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATG 355
                | ||| |        |  |   |   ||    | |  || |  |||||||||||||||||
Db         1118 AGTGCAACCTTACCTTTTTTCACTTCAATGGCAGCACCTAAGCCATCAATTGCTGCAAGT 1059

Qy          356 GTTGAGCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCA 415
                || |  || |  || |||||||| || |||||| ||||         ||  |     | |
Db         1058 GTAGCACCAACATTAGTACCTTTAGAGTCATTATAATAACGGACATCATGCACGGTTTTA 999

Qy          416 ATATATTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCC 475
                |  || ||||||||| |||  ||| ||||| ||  | || |  || |||||||| |  |||
Db          998 ACGTACTCACAGCGGTGCTCTAGTCCTTTAAATTTTTTTAAAGTTTCAAGCATTGACTCC 939

Qy          476 ATGGGAAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGC 535
                || || | |      |||    || || || |    |  ||| | |||||   ||  || |||
Db          938 ATAGGTAAACCAATTGCTTCCCCCAATGCTAAACAAGCTAAAGCATTAGCTACATTGTGC 879

Qy          536 TTGCCTTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCC 595
                 | ||||  || | ||| |    | ||   |  |     | |     | ||   |
Db          878 ATACCTTGAATATATAAATCTGAGCTTTTAATTAACCGCTGTAAACCACGAGCAAGCCAA 819

Qy          596 ATACCATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATTT 655
                |    | |  |  ||||   ||||     || |    |        |||  |  | |   |
Db          818 AGCGTACCGTCGGCATCTCTTAAAACGCCATATTGATT--------TAAATCCGGTGCGT 767

Qy          656 GTTGTGATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCT 715
                 |      ||   |   ||| | | |    || |     |||| |
Db          766 TTAAACCAAAGCTTTGCATTGGTGTCGTATCTGGAACAAGTGGACGGCTTA--------- 716

Qy          716 AAATTGGCATCGTCTTGGTAGATGATGGCGGTTTGGCAATTTTCAAAAATACGCAATTTT 775
                      ||||| ||| |||  |  |   |  |||        ||  ||||||||   ||||
Db          715 ----ACGCATCATCTCGGTTAAATACAACTTTTTTAGCGCCTTGGAAAATACGATGTTTT 660

Qy          776 TGTGCCAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAA 835
                  |     |||  ||| |||     | ||   || || | ||| ||  |     | ||| | |
Db          659 GCTTGATGATACCCCAGCATATTTCCATGACGGTCTAAATGGTCTTCACTCATATTGAGA 600

Qy          836 ATTGTTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCA 895
                |    |||  | | |||    || |  |||  |    || | |||||| ||||| | ||||
Db          599 ACCACTGCTACCTCAGCGTTTAAGTGAGAGGTGGTTTCTAACTGAAAGCTTGATAACTCA 540

Qy          896 AGCACCACCAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTA 955
                || ||||  || ||      |||     | | ||| || |   ||  ||    ||||  ||
Db          539 AGGACCAGTAACTCTGG---TTGATCTTTAAGTAAATCTAACGCTGGTCGACCAAGGTTG 483

Qy          956 CCGCCGACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACA 1015
                ||||| ||  ||||  || | |||||||  || |||||    ||| | ||| || || |||
Db          482 CCGCCTACTGCAACTTTCTTACCTGCATCCTTAGCCATTAAGCCAATTAAAGTGGTTACA 423

Qy         1016 GTGCTTTTGGCATTTGAGCCTGTGATGGCGATGATTGGTGTGTTTGTTGCTTGAGTTAGA 1075
                || ||||| ||||| || ||||| || || |  || ||    |  || ||
Db          422 GTACTTTTTGCATTGGAACCTGTAATCGCCACAATCGGCACATCAGTGGC---------- 373

Qy         1076 TCTCGTGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTGA 1135
                                 |      || ||  || | |||||   | ||||| ||| |||| |
Db          372 -----------------ACGGCGCAATAACTGGATATCGCCCACCACAGAAATACCTTTA 330

Qy         1136 GCTTTGGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGA 1195
                ||  | || ||   |||   |||   |  |||   || ||| ||||| |  || ||||
```

Db   329 GCAATAGCTGCCTGAAGTTCCGGTGAATTGTGGGTGCAAGGCCTGGGGCTTAAAATAATCTCT 270
Qy   1196 TCGGCTGAGCTTAAGAGTTCACTGTCCAAAAATGACAATGAACGCCTTGGGGC 1255
Db   269 TCTCGTTTGTAATATAATTCTTGATCAAGCTGACCAAAAACTGGTTTTAACCAGCCCGGA 210
Qy   1256 AGTTTATCTGCCAAGGGTAGGATTGGGATTACCCATGCATCGATGATGCTAAACTGGTTTTGTGACCGGTGA 1315
Db   209 ATCTGATCGTGTCCGGGAGGGTGTGGGGCCGGGGAAATCCGTTACAGCAACTTGGTTAGCCTTGT 150
Qy   1316 TTGACCAAAAAATTCACACTGCAGTCCCAGAGCGTGCCCAAAACCGACGACGGCAATAT 1372
Db   149 TCATGCAGGAGAAAATTTCACAGCAGAAAACACCTGATATTCCCAAAGCCTGCTACAACTTTT 93
```

```
Query Match    11.6%;  Score 160.4;  DB 3;  Length 1830121;
Best Local Similarity  49.6%;  Pred. No. 9.6e-37;
Matches  683;  Conservative  0;  Mismatches 606;  Indels  89;  Gaps  7;
```

```
Qy      1  TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATATTAAATTGATC  60
           |||      |||||||||||||||||||||||   |||||||| || |||||||||
Db 1204439  TAAACATTGAGCTAAAATGCCTAAAATTCTTCGCCCGGGGTGCTTTCAAAAGAAGCAAAACTGATC  1204380
Qy     61  CATGCTAGCACACAGGGTGACAACACAGCCCTAGCGGGTGCGTTTGACTGGCTGCCAA  120
           || |||  || ||||| ||  |||||   ||       |||  ||||  ||||  |||
Db 1204379  GAGACTTGCACAAGCAGGCGACAATAATACCA--------TATCTCCGCTTTGCAA  1204332
Qy    121  TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGATGCCAAAATGTGCATCACCTTT  180
           |||                             |||||||||||||||||||||||||||||||
Db 1204332  ---------------------------------TATCTCCGCTTTGCAA  1204332
```

Alignment continuation (right column):

```
Db 1204331  TGTTGGGGCGTAAAAATTCTATCGCTTGTTCCATTGTATCGGAACAA------ATAACTTTG  1204278
Qy    361  GCGGATGTTTGTTGTACCTTTTTGAATTCATTAAAATATGCCTTGCTTATCAATATCAATATA  420
Db 1204277  CGATGAAAATTTTGGCAAGCAGCGGCACCATCTCGACGATTAAAATACCACTATCCACCA  1204218
Qy    421  CGATGAAAATTTTGGCAAGCAGCGGCACCATCTCGACGATTAAAATACCACTATCCACCA  480
Db 1204217  TTGATTAATTAATTCAGCTAAATTCTGACAAAAAATCAGCACCCTTTTCCGTCGGCTTAGCAA  1204158
Qy    481  AAGATTGGCACGCTGTCCTAAGGCAGTGCCAGATAAGGCATTGAGTGAGTGAAGGTTATGCTTGCC  540
Db 1204157  CAAATGCAATTTACCCTCAATATA---------AAGCCAGCAAATGCAGCAACTGTACT  1204107
Qy    541  TTTGATTTTTAATTGATGGATTGGGTAAAAGTGTTGTTTTTTTGATGATAATATGCCATACC  600
Db 1204106  CCCCACATTTTTGTGCTTTAGAGTCATTAATCCAACATTCATTAATAGTCGACC  1204047
Qy    601  ATCACCCATCATCAAATAAATAAAATCTCG--CGTTTGGGTGGCTAAGTAAGCTATTTGTTG  659
Db 1203986  AATACCTATAGCTTGTGCCAATGCGTTCTCAGTAATTCCGCAAAAAGAAACGGGTATGTTT  1203927
Qy    660  TGATAATGGTGTGGATTGGGGGCAACGCTTGTCAGTGTGGTTCAAGCATTGCTTGGCTAAAT  719
Db 1203926  AACCAAATGTAGCTTCTTCACAAGGTAAAATCACTTCATCTTTTACCATTATAATATGCTTT  1203867
Qy    720  TGGCATCGTCTTGGGTTAGAGATGAGGCGGTTGGCAATTTCAAAATACGCAATTTTTGTG  779
Db 1203866  GCCAATTTTCAGTTTTTAGCCATTAATCCGGACTATTTCGCAAAAGGGTATGTTTTTGCTT  1203807
Qy    780  CCAAATAAATCCACCATGCCGGATCGTTCAAGAGATGATCAGCCGGAGAGATTTAAAATTG  839
Db 1203806  CGCTTGATTTTCGTTTTCCCCAAAAGTCAGCC---------  1203775
Qy    840  TTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA  899
Db 1203717  GGCGGATAATCTTCTAAATCCATATAGCGATCCCATATGATCTTCAGTCGTTCAAGACAG  1203658
Qy    900  CCACCAAAATCCATATTTGGAATCGTTAATAATTCAAGTGCAGGGGTGCCAAATATTACCGC  959
Db 1203657  TCGCTCCCGCAGCCTTTTAACGTATAAAGTTGTCTCAAGCGTAAAAACTAGAAAAGCTCTAGTA  1203598
Qy    960  CGACACCAACAAATACCTGCATGTTTTGCCATCTCGCCTTGTGTTGTTTGTTGCTTGAGATCTC  1019
Db 1203597  CATAAAAGTTCACAATCTTTCATTC---AACAATGACAAAGCGGGAATCCCAATAATTCCCGC  1203541
Qy   1020  TTTTTGGCATTTGAGCCGTGTGATGCGGATGAGTTGGTGTTGCTTGAGTTAGAGATCTC  1079
Db 1203540  CCATACCAACAAACTTCACCAGCAGCTTCGCCGGTTTTCCATCTCCGAGTGGTGAGTGGTAC  1203481
Qy   1080  GTGCTTGAGTGTATCAAATGAACATCGCTGATCACAGGAATGCCTTGAGCTT  1139
Db 1203480  TTTTTACCATTTGAACCTGTAATCCCACCAATTGTCGCTCGCGCGCAGA------  1203429
Qy   1140  TGGCGGCGACAATACTCGGCCGGTTCGGGGACGATTACTTCCACTCCGCTTAA  1199
Db 1203428  -------------------GAATAATTCAATATATGGCGATTACTAACATCGGTAGATCGG  1203388
Qy   1200  CTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAAATGAACGCCTTGGGCAGTTT  1259
Db 1203387  GTGCGGTTTGAATTTTCTGGTGTTTTTACCCCAAGCCCTGGGGCTAATACATCATATCGC  1203328
```

```
Db 1203327 TTTCAAGTAACCATTCCTGATTAAACTACCAGTATGATGAAGAGGAATTATTTTGAGGAAGTT 1203268
Qy 1260 TATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGTGACCGTGATTGA 1319
Db 1203267 TATCAATACCAGTAGGGATTTTTTCCGAGTATCAATCACACGAATATTAGCCTTGTGTGGGATA 1203208
Qy 1320 CCAAAAAATTCACAGCCTGACAGTCGCCCAAAACCGGACGACGGAATGGAATGTTTTTTTT 1377
Db 1203207 AGAGATAATCCACACAAGAAAGAGACCCTGTTTTGCCAAGCCCGATGATGAGTAATATATTTTT 1203150
```

```
Qy    1 TAAAGCGTGAACCATGTTAACAACTTATCACCTCGCTCATTATTATAATTTTTAAATTGATC 60
Db 1204439 TAAACATTGAGCTAAATGCGTAAATTCTTCGCCGCGCCTTTTCAAAGAAGCAAACTGATC 1204380

Qy   61 CATGTCTTAGCCACAAGCAGCAGTGACAAAAGCACAGCGCCCTAGCGGTGCTTTTGACTGGCTGCCAA 120
Db 1204379 GAGACTTGCACAAGCAGGCGCACAAATAATACCA-----------TATCTCCGCTTGCAA 1204332

Qy  121 TACTGCCCGATCAGTCGTACCAACCTGATATATTTTTAGATCGCATCACCTTT 180
Db 1204331 TGTTGGGCGTGAAAAATTCTATCGCTTGTTCCATTGTATCGAACACA------ATTACTTTG 1204278

Qy  181 GAGTAAATCTTGCTCAAATTATCGGGGGCATCTTCTTGACCGATTGACACAACTATTCCACCA 240
Db 1204277 CGATGAAAATTTGCAAGCAGCCACCCATCTCGACCAAAAACAATAACCATATAATGTGG 1204218

Qy  241 CTGCATAACATATGCTTGACAGCTCACTAAAAATCTTGACCTTTTTACCAAAGCCCACCCAAAAT 300
Db 1204217 TTGAATAATTAATTCAGCTAAATTCTGAAAAATCAGCGCCCCTTTTCCGTCTCCGCTAGCAA 1204158

Qy  301 CAAATGCCAGTGATGATATGCCTAAAACCGTGCCAAGCCCATCATCGCTTCCAAATGGTTGA 360
Db 1204157 CAAATGCAAATTACCCTCAATATA-------------AAGCCCAGCCAAATGCGCAAACTGTACT 1204107

Qy  361 GCCGATGTTTCTACCTTTGTGATTCATTAAAAAATATGCCTTGCTATCAATATCATCAATATA 420
Db 1204106 CCCCACATTGTTGCTTTAGGTGTCATTAAATCCAAGCACGGAAATGCCATTGAGTGACTAA 1204047

Qy  421 TTTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATTCAAGCATGGCATCCATTGGG 480
Db 1204046 TTGAAAACGATGATCTAACCCTTTGAAATGACGAACGTGCCGGTTACCGAGTTGAATCTAAATT 1203987

Qy  481 AAGATTGGCACGCTGTGCCTAAGGCAAGTGCAGAATAAGGCATTGAGTAGGTTATTGCTTGTG 540
Db 1203986 AATACCTATAGCTTGTGCCAAATGCTGTGTGCCAAAAAGTTCATTATTATGGGACC 1203927

Qy  541 TTTTGATTTTTAATTGATTGGATTGGGGGGCAACCGTTGTCAGTGGTGGTCAAGCATTGCTTGGCTAAAT 600
Db 1203926 AACCAATGTAGCTTCTTCACAAGGTGAAAATCACTTCATCTTTTGTTTTTGATGATAATAATATGCTT 1203867

Qy  601 ATCACCATCATCAAATAAAATAAAATCTGC--CGTTTGGTGGCCTAAGTAAGCTCATTTTGTTG 659
Db 1203866 GCCATTTTTCAGTTTTTAGCCAAATAAATCCGCACTACTTTTTTCCCGCAAAAAAGAAACTGATGTTT 1203807

Qy  660 TGATAATGGTGTGATTGAGTTGGGGGGCAACCGTTGTCAGTGGTGGTCAAGCATTGCTTGGGCTAAAT 719
Db 1203806 CGCTTTGAATTTGTTGTTTCCCGGATAAGGTCAGCC---------- 1203775

Qy  720 TGGCATCGTCTTGGTAGATGATTGGGCACGTTTGGCAATTTCAAAAATACGCAAATTTTGTG 779
Db 1203774 ---TATCTTCATTGTTCAAACACACTTTAGCATTATGATAAAATGCGTAATTTTTGCTT 1203718

Qy  780 CCAAAATAATCCACCATGCCATCGTCGCGATCAAGATGATGATCAGCGGAGAGATTTAAAATTG 839
Db 1203717 GGCGATAATCTTCTTAAATACATTATACGATCCATAATCTTCAGTCACGTTCAAGACAG 1203658

Qy  840 TTGCCCCTTGAGCTGCCAAATTTGAGATAATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA 899
Db 1203657 TCGGCTGCCGCAGCTTTTAAGCTTATAAGTGTCTCAAGCGTGAAAACTAGAAAAGCTCAGTA 1203598

Qy  900 CCACCAAATCCATTCATTTTGGATCGTTGATTAATTCAAGTGCCAGGCGGTGCCAATATTACCGC 959
Db 1203597 CATAAAGTTCACAAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAAATATTTCCGC 1203541

Qy  960 CGACACCAACAAATCATGCCTGCACTGTTTTCCCCATCCTACTAATGTGTTAACAGTGC 1019
Db 1203540 CCATACCAACTTTCACACACCAGCCTTTCGCCATTCATAAAACTAAAGATTACCGGTAC 1203481

Qy 1020 TTTTGGCCATTTGAGCCGTGATGGGCGGATGATAGTTGGTGTTGTTTGCTTGAGTTAGATCTC 1079
Db 1203480 TTTTACCATTTGAACCTGTAATCCCACACATTGGCTTTTGTGTCGCGCACAGCA-------- 1203429

Qy 1080 GTGCTTTGAGTGTATCAATGAACATCGCTGATCACAGGAAATGCCTTGAGCTT 1139
Db 1203428 --------GAATAATTCAATATCGGGGGTTCTTGGGGTTAATCGGCGATTACTTCCACTCCCGATCGG 1203388

Qy 1140 TGGGCCGGCGACAATACTCGGGCGGTTTTCTTGGGTTAATCGGCGATGATGATCTCATTACGG 1199
Db 1203387 GTGCGGTTTGAATTTCTGGTGTTTTTTACCGGCAAGCCCTGGGCTAATAACAATCATTACGC 1203328
```

```
Qy    1200  CTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACANATGAACGCCTGTGTGGGCAGTT 1259
                  |||||||||||||||||||||||||||||||| ||||||| ||||  ||||||| |||
Db  1203327  TTTCAAGTAACCATTCCTGATTTAAAACTACCAGTTGACGATGAGAGGGATATTTTGAGGAAGTT 1203268

Qy    1260  TATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGA 1319
                  |||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||
Db  1203267  TATCAATACCAGTAGGATTTTTTCGAGTATCAAATCACACGAATATTAGCCTGTGTGGGATA 1203208

Qy    1320  CCAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAAACCGACGACGGGCATTATTTTTTT 1377
                  |||||||||||||||| |||||||||||||||||||||||||| |||||||| |||||
Db  1203207  AGAGATAATCCACACAAGAAAAGAACCTGTTTTTGCCAAGCCCGATGATAGTTAATATATTTTT 1203150
```

RESULT 6

US-10-158-865-1/c
; Sequence 1, Application US/10158865
; Patent No. 6846651
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; Patent No. 6846651
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2C1D1
; CURRENT APPLICATION NUMBER: US/10/158,865
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/426,787
; PRIOR FILING DATE: 1995-06-07
; PRIOR FILING DATE: 1995-04-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 1
; LENGTH: 1830121
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (4747)..(4747)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (9921)..(9921)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10150)..(10150)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (29298)..(29298)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36543)..(36543)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36551)..(36551)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (36636)..(36636)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (40808)..(40810)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44416)..(44416)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44905)..(44905)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (44975)..(44975)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45732)..(45732)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45593)..(45593)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51334)..(51334)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (47036)..(47036)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51602)..(51602)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51786)..(51786)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (51805)..(51805)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55369)..(55369)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65309)..(65309)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (65313)..(65313)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (80024)..(80024)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (100091)..(100091)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (102696)..(102696)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (105121)..(105121)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (107248)..(107248)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (117136)..(117136)
; OTHER INFORMATION: n equals a,t,c, or g
; FEATURE:

```
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
```

```
Qy         1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTGCTCATTATAATTTTAAATTGATC 60
                 |||     ||       |          |       |   |        |||||
Db   1204439 TAAACATTGAGCTAAATGCGTAAATTCTTCGCCGGCTTTTCAAAAGAGAAACTGATC 1204380

Qy        61 CATGCTTAGCACAAGCACAGGTGACAAAGCACAGCCCCTAGCGCGGTGCTTTGACTG 120
                    | |||||||||  | | ||||||||   |                |||| |||
Db   1204379 GAGACTTGCACAAGCAGGCGGACAACAATAATACCA-----------TATCTCCGCTTTTGCAA 1204332

Qy       121 TGTTGGGCGTGTAAAAATTCTATCGCTTGTTCCATTGTATCGGAACCA 180
                   ||||||||                            ||||||
Db   1204331 CGATGAAAATTTTGCAAGCAGGCGACACCATCTGACATATATTTTTAGATGCGCATCACCTTT 1204278

Qy       181 GAGTAAATCTTGCTCACAATTATCGGGGCATCTTGATCGGTGGTAAATAACACTCACCA 240
                | |||||  | |||| |     ||||| |  | ||||| || |||
Db   1204277 CGATGAAAAATTTTGCAAGCAGGCGCACCATCTGACTCTTGGATATAATTTTTTAGATGCGTGG 1204218

Qy       241 CTGCATAACATATGCTGACAGCTCACTACAAAATCTTGACGTTATTATTTTTGAGATCCAAAAT 300
                |||||  |  ||     ||| | || | |   | ||||| |   |  |  |  |  |||||
Db   1204217 TTGATTAATTAACTTCAGCTAATTCTGAAAAATCAGCCCCTTTTTCCGTCTCCGCCCTAGCAA 1204158

Qy       301 CAATGCCCAGTGATGATATGCCATAAAACCGTGCCCAAGCCCCATCAAATTGCTCAAATGGTTGA 360
                ||  | ||  ||   |  |  |  |  | |   |    |||||   || |  | || ||| |
Db   1204157 CAAATGCAATTACCCTCAATATA---------AAGCCCAGCCAATGCAGCGAACTGTACT 1204107

Qy       361 GCCGATGTTTGTACCTTTTGAATCATTAAAATAATGCCTTGGCTATCATAATCATCAATATA 420
                ||| |  ||| |  |||||||| ||     ||  | |     || |   || |    |  |
Db   1204106 CCCCACATTTGTTGCTTTAGAGTCATTAACCAATGTCCATTTTAAGCTGAAATATTTTTCGC 1204047

Qy       421 TTCACAGCGATGCGGGTAGCCCCTTTGAAAGTCTTAAGGGGTATCAAGCATCCATGGG 480
                ||   |||| |  | |  |   || ||| |||  || ||||    |    |||  |||
Db   1204046 TTGAAAACGATGAATCTAACCCTTTGAAATGAACGAAGTGCGGGTTGCAATCTAAATT 1203987

Qy       481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAAGACGATTGACTAGGTTTATGCTTGCC 540
                |||   |||  | | || |  | | ||  | || |  ||  |||||  ||| |  ||| ||
Db   1203986 AATACCTATAGCTTGTGCCAATGCGTGTTGCCCAAAATGGTTCATATAATTTAGGGGACC 1203927

Qy       541 TTTGATTTTTAAATTGATGGATTGGTAAAAGTTGTTTTTTTTTTTGATGATATAATGCCATACC 600
                 |||||   || ||   ||||| |  | ||  |  |  ||      |||| |   ||| |||
Db   1203926 AACCAAATGTAGCTTCTTCACAAGGTAAAAATCAGCCGTTTTTACCATTATAAAAATATGCTT 1203867

Qy       601 ATCACCCATCATCAAAATAAAAATCGC-CGTTTGGTGGCCTAAGTAAGCTATTTGTTG 659
                |   |||| |  |||| |   |   |    |  ||| |   ||||   ||   ||||
Db   1203866 GCCATTTTCAGTTTTTAGCCAATAATCCGGCAAATAATCCGGCAAAAGAAACGGGTATGTTT 1203807

Qy       660 TGATAAATGGTGTGATTTGGGGCACGCGTTGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAAT 719
                 |||   | ||   ||  |||  | | |  |   | | ||  |  |  |||   ||||| |||
Db   1203774 ---TATCTTCATTGTTCACACACTACTTTAGCATTATGATTAAAATGCGTAAATTTTGCTT 1203718

Qy       720 TGGCATCGTCTTGGTGTAGAATGGATGATGGGGTTTGGGCAATTTTCAAAAAATACGGCAATTTTGTG 779
                 ||| | | |  |||  ||  | || | || |||||| | ||  |||   | |   | || | ||
Db   1203717 GGGCGGATAAATCTTCTAAATCCATATAGCGGATCCAATATGATCACGTTTTCAAGACAG 1203658

Qy       780 CCAAATAATCCACCATGCCGATCAAGATGAGATGCAGCGGAGAGATTTAAAATTG 839
                |||||  |||| | ||    || || ||     |  | | ||| |  ||   | ||
Db   1203657 TCGCTCCCGCGACAGCGTTTTAAGCTATAAGTGTCTCAAGCTGAAAACTAGAAAGCTCTAGTA 1203598

Qy       840 TTGCCCCCTGAGCGTGCCAAATTGGGTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA 899
                 |||   | | |   || ||||  |  |  | | ||  | ||||      || | |  ||||
Db   1203597 CATAAAGTTCACAACTTCAATTC---AACAATGACAAAGCGGGAATCCCAATATTTCCGC 1203541

Qy       900 CCACCAAAATCCATATTTTGGATCGTTAAGTAGTGGAGGCGTGCCAATTATTACCGC 959
                |||  ||| |||| | ||| | || ||  |  |  | || |  || ||||| |||||
Db   1203540 CCATACCAACTTCACCAGCAGCTTTCGCCCATTCAATTCAATAAACTAAAGTAGTTACGGTAC 1203481

Qy       960 CGACACCAACAAATCCATATCTTTGCCATCCTGCCTACTACTGGTTAATGTCGTAACAGTGC 1019
                ||| ||| ||   | | || ||  ||| ||    | | |  ||  |  ||    | || |
Db   1203480 TTTTACCATTTGAACCTGTAATCCCACAACAATTGGCTTTTGTCGCTCTGCGCGGCA 1203429

Qy      1020 TTTTGGCATTTGAGGCCTTGTGATGGCGGATGATTGGTGTGTTTGCTTTGAGTTAGATCTC 1079
                |||||||||||||||       |    ||| || | | |   | | ||     |||  |||
Db   1203480 TTTTACCATTTGAACCTGTAATCCCACAACAATTGGCTTTTGTCGCTCTGCGCGGCA 1203429
```

```
Query Match  11.6%;  Score 160.4;  DB 3;  Length 1830121;
Best Local Similarity  49.6%;  Pred. No. 9.6e-37;
Matches 683;  Conservative 0;  Mismatches 606;  Indels 89;  Gaps 7;
```

```
RESULT 7
US-09-543-681A-2830/c
Sequence 2830, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABI[L]
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
SEQ ID NO 2830
LENGTH: 1323
TYPE: DNA
ORGANISM: Proteus mirabilis
US-09-543-681A-2830

Query Match      9.1%;   Score 125.2;  DB 3;  Length 1323;
Best Local Similarity   53.5%;  Pred. No. 1.1e-27;
Matches  352;  Conservative   0;  Mismatches 273;  Indels 33;  Gaps 3;
```

```
Qy        1080 GTGCTTGAGTGTATCAATGAACATGCGTCTGATCTCACGGAATGCCTTGAGCTT 1139
Db     1203428 ---------------------------------GAATAATTCAATATGCGATTACTTCACTCCCGCCTTTAA 1203388
Qy        1140 TGGCGGCGGACAATACTCGGCGGTTCTTGTGGGTTAATGCCGGGGCGGGGCTGATGATGACTCGG 1199
Db     1203387 GTGCGGGTTTGAATTTCGTGGTTTTACCGCAGCCTGACAAGCCTTGGGCTTAATAACAATCATATCGGC 1203328
Qy        1200 CTGAGCTTAAGAGTTCACTGTCCAAAATGACAATGAACGGCCTGTGGGCAGTT 1259
Db     1203387 CATCATTAAGCCAATACTTGATGAAAAACTCCCCGTAATACAATCATATCGGC 1203328
Qy        1260 TATCTGCCAAGGTAGGAATTGGGGATTAACGCTGGGGAAGTACTCATCGATGATGTTGTGACGTGATTGA 1319
Db     1203327 TTTCAAGTAACCATTCCTGATTTAAACTAACTTGAACGAATATTTTGACCCTGATTGA 1203268
Qy        1320 CCAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAAACCGACGACGGCGCATATTTTTT 1377
Db     1203267 TATCAATACCAGTAGGAATTTTTGAGTATCAATCACACGAATATTAGCCTGTTGGGATA 1203208
Db     1203207 AGAGATAATCCACACAAGAAAGACCTGTTTTTGCCAAGCCCGATGATAGTAATATATTTT 1203150
```

```
RESULT 8
US-09-252-991A-7702/c
Sequence 7702, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7702
LENGTH: 1371
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7702

Query Match      8.5%;   Score 117;  DB 3;  Length 1371;
Best Local Similarity   48.4%;  Pred. No. 3.7e-25;
Matches  506;  Conservative   0;  Mismatches 500;  Indels 39;  Gaps 5;
```

```
Qy          17 TTAACAAACTTATCACCTCGCTCAATTATAATTTTTAAATTGATCCATGCTAGCACAAGCA 76
Db        1349 TTGGCGAACAGGCGTCCGCGCGTTCTTCGAAGTTCTTGAACATGTCAAGGCTCGCGGAGGC 1290
Qy          77 GGTGACAAAAGCACAGCCCTTAGGCGTGCTTGACTGGCGTGCCAATACTGCCCGATCAAGT 136
Db        1289 GGGCGACAACAGCACCGCACCATGCCCTTCGC---------GGGCCAGTCTCGGCGCTGCCGG 1239
Qy         137 GTACCAAACCTGATATATTTTAGATGCCAAAATGTGCATCACTTTGAGTAAAAATCTGCTCA 196
Db        1238 ACTGCTTCGTCCCAGCGGTTGGCGACGCCACCAGGCTACCGCCTTGCCCAGTGCCTGGGCA 1179
Qy         197 ATTATCGGGCGATCTTGACCGGCCAATTAAATACACAGCTCAATAACATATGCT 256
Db        1178 ATCAGCCCGCGGTCACGCGCGAAGCAGTACCACCTGCGGTCACCCGGCTCG 1119
Qy         257 GACAGCTCACTCAAAAATCTTGACGCTGTTACCAAGCCCACCCAAAATCAAATGCAGTGATGAT 316
Db        1118 CGCAGGTCATCGAGAAATCGGCGCCCGTCCGGCGGGAGCCGAGCTTTGGTACCT 1059
Qy         317 ATGCCATAAAACGTGCCGAAGCCCATCAATGCTGCAAATGGTTGAGGCGAGTTTTGTTACCT 376
Db        1058 TCGATGT---CGGCACCCGCGCCTTGCTTGTATCAATATCATCAATTTCACGCGGGT 1002
Qy         377 TTTGAATCATTAAAATATATGCCTTGCTATCAATATATCCGGCGCCCCGCGGCGTTGGGGC 436
Db        1001 TTGGAATCGTCGTAGTCTCACGCGCGCTGGGACCCACGGGACTCACGCGGATGAGCC 942
Qy        1080 GTGCTTGAGTGTATCAATGAACATTGACCGACAACAATTTCCACTCCCATGGATCTC 1139
Db         311 --------------------------CAAAATAGCTCAATATCACCGACAACAATTTCCACTCCCATGATCTC 269
Qy         437 AGCCCTTTGAAAGTCTTAAGGGTATCAAGACATGGCATCCATGGGGAAGAATTGGGCACGCTGTG 496
```

```
RESULT 9
US-09-252-991A-7861
; Sequence 7861, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
;  APPLICANT: Marc J. Rubenfield et al.
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
;  TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 107196.136
;  CURRENT APPLICATION NUMBER: US/09/252,991A
;  CURRENT FILING DATE: 1999-02-18
;  PRIOR APPLICATION NUMBER: US 60/074,788
;  PRIOR FILING DATE: 1998-02-18
;  PRIOR APPLICATION NUMBER: US 60/094,190
;  PRIOR FILING DATE: 1998-07-27
;  NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7861
;  LENGTH: 1401
;  TYPE: DNA
;  ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7861

  Query Match             8.5%;  Score 117;  DB 3;  Length 1401;
  Best Local Similarity   48.4%;
  Matches  506;  Conservative    0;  Mismatches  500;  Indels   39;  Gaps    5;
```

```
Qy     941 AGGGCCGGAAAACGCCTTCAGCGCGCCGAGCGTCGAACGGGCGCAGGCGCCGACCCCATG 882
           ||||  |  ||  ||    |    |    |   |    |   |   |   |  |  | |
Db     497 CCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCCTTTGATTTTTAATTGA 556

Qy     881 CCCAGCCGCAGCCGGGCGAGCGCGTTGGAATAGTTGTGGGCGCGACGATCTTCAGTTCG 822
           |  |  |  | ||  || |  ||  | | |      |   |    | ||  ||  | |
Db     557 TGGATTGGTAAAAGTTTGTTTTTTTGATGATAATAATGCCATACCATCACGATCAATT 616

Qy     821 CCAACCGGCAGCAGCTTGTTCGGAACCGACCACTTCTGGGGTGTCTTCCCTCGATC 762
           |  |  | | ||  ||  |   |   |  |  ||    |   |    |   ||
Db     617 AAATAAAAATCTCCGTTTGGTGGCTAAGTAAGCTATTTCGTTGTTGATGAATGGTGTGAATTT 676

Qy     761 AGGCCGAAAGC------------CTTGAAGTCCGGCGCTTGTTCAGCCGGAA 724
            || ||  |  |            || ||  | | |  ||  | |  | |
Db     677 GGGGCCAACGCTTGTCAGTGTGGGTCAAGCATTGCATCGTCTTGGTAG 736

Qy     737 ATGATGGGCGGTTTGGCAATTTTCAAAAAATACGCACCATG 796
           |    | |   |    | |     |  ||    | | | ||
Db     677 GACAGCTCACTAAAAGAATCTTGACCTTTACCAAGCGCTTT 606

Qy     677 GGGGCCAACGCTTGTCAGTGTGGTCAAGCATTGCATCGTCTTGGTAG 736
           |  | |    |     |  |  |   |  |   | |  |    |
Db     665 ACCACGACCTGCGCGGGCCACCGGGGAGAGATCCGGTCG 606

Qy     605 CCGTCGTACGGATCCATATGGTCTTCGCTGACGGTTCAGCACGGTTG 546
           |   |   |  |  | ||  |    |  | |   |  | |  |
Db     737 ATGATGGGCGGTTTGGCAATTTTCAAAAAAT 796

Qy     545 AGGCGGATCGCCAGTTTTCCAGCTCGGAAGCTCGACAGCTCCAACAGTACGACGTCG 486
           |   |  | |  |       | |    |     |   |       |  |    |  |
Db     797 CCATCGGTGCCGGATCAAGAGATGATCACGCGGGGAGGAGAGGATTTAAAATTGTTGCCC 856

Qy     485 TCG---GCCAGCAGGTCGAGCCGCCGGGGTGCCCAATATTACCGCGGCACCAACAATCATG 429
           |      |  | || | |  |   | |   |   |  | | | |   |  |   | |  |
Db     857 AAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCA 916

Qy     917 TGGATCGTTAATAATTCAAGTGCAGGCGGTGCCCAATATTACCGCGGCACCAACAATCATG 976
           |    | |  |  |    |  |  | |  |    |  |  |      |
Db     917 TGGATCGTTAATAATTCAAGTGCAGGCGGTG 976

Qy     857 AAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCAAAATCCATATTT 916
           |       |   |  |   |  |  |  | |   |  ||  |  |    |  |
Db     377 TTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATATTCACAGGATGG 436

Qy     797 CCATCGGTGCCGGATCAAGAGATGATCACGCGGGGAGGAGAGGATTTAAAATTGTTGCCCTTGCC 856
           |   |  |  | | |     | | |  |  | |    |  |  |   |   |  |  |  |
Db     362 TTGGAATCGTGTTAGTAGCTCACGGGTACCCACTGGCCAGGGATGAGCC 421

Qy     737 ATGATGGGCGGTTTGGCAATTTTCAAAAAATACGCACCATG 796
           |    | |   |    | |     |  ||    | | | ||
Db     317 ATGCCATAAACCGGTGCCAAGCCCATCAAATTGCCGTCTTCCCGGCGTCCGGCGATT 376

Qy     677 GACAGCTCACTAAAAGAATCTTGACCTTTACCAAGCGCTTT 606
Db     257 CGCAGGTCATGGGAAAATGGGGCGGCCCCTTTGTCCGGTCA--GGGCATCGGCCGGATTC 304
```

```
Qy    1037 GTGATGGCGGATGAATTGGGTGTGTTTG 1061
           ||||||||||||||||||||||||||
Db     368 GTGATGGCGGACGGATCGGGCGGCCTTG 344

RESULT 10
US-09-701-229-1/c
; Sequence 1, Application US/0701229
; Patent No. 6890910
```

`,`

GENERAL INFORMATION:
; APPLICANT:
; APPLICANT: El-Sherbeini, Mohammed
; APPLICANT: Azzolina, Barbara
; TITLE OF INVENTION: MURD PROTEIN AND GENE OF PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA
; FILE REFERENCE: 20193P
; CURRENT APPLICATION NUMBER: US/09/701,229
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/087,308
; PRIOR FILING DATE: 1998-05-26
; PRIOR APPLICATION NUMBER: PCT/US99/11585
; PRIOR FILING DATE: 1999-05-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-701-229-1

Query Match   8.5%; Score 117; DB 3; Length 1450;
Best Local Similarity   48.4%; Pred. No. 3.8e-25;
Matches 506; Conservative   0; Mismatches 500; Indels 39; Gaps   5;

Qy   17 TTAACAAACTTATCACCTTGCTCATTATATAATTTTTTAAATTGATTCCATGCTAGCACAAGCA 76
Db   1375 TTGGCGAACAGGCGGTCCGCGTTCTTCTTCGAAGTTCTTGAACATGTCCAGGCTTGCGCGCAAGT 1316

Qy   77 GGTGACACAAAAGCACAGCCCTTAGCGGTGCTTGCTCAATACTGCCGGATCAAGT 136
Db   1315 GGGGCACAACAGCACCGCATCGCCTTGCTTGACTGGCTGCCCAATACTGCCGGATCAAGT 1265

Qy   137 GTACCAACCTGATATATTTTTTAGATGCCAAATGTGCATCACTTTGAGTAAATTCTTGCTCA 196
Db   1264 ACTGCTTCGTCCAGCGTTGCGACCGCGCACCAGCGTTGCCCAGTGCCTTGGGCA 1205

Qy   197 ATTATCGGGCGCATCTTTGACCGGATTAAATACACACTATTAAACAATATGCT 256
Db   1204 ATCAGCCCGGCGTCACGGCCAAGCAGTACCACCCCACTGCATTAAACATATGCT 1145

Qy   257 GACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAATCAATGCCAGTGATGAT 316
Db   1144 CGCAGGTCATGGAAAATCGGCGCGCCCCTTTGCCGCTCTCCGCGAGCAGCACCAGCTTGCCG 1085

Qy   317 ATGCCATAAACCGTGCCAAGCCCATCAAATGGTTGAGCGGATGTTTGTTACCT 376
Db   1084 TCGGATGT---CGGCACCCCCGATCGCCCAGGGCGCCGCAGCGGTTGGTTGGCC 1028

Qy   377 TTTGAATCATTAAAAATATGCCTTGCTTGTCTATCAATATATTCACAGCGATGCGGT 436
Db   1027 TTGGAATCGTCGTAGTAGCTCACGCCTGCCGTACCACTGGCGGATGAGCC 968

Qy   437 AGCCCTTTGAAAGTCTTAGGGTAATCAAGCATCCATGGGAAGATTGGCACGCTGTG 496
Db   967 AGGCCGGAAAAACGCCTTCAGCGGCGTCGAACGGGAGCCGAGCCGCCATGG 908

Qy   497 CCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATAGCCTTGATTTTTAATTGA 556
Db   907 CCCAGCGCCAGCCCGGTTGGAATAGTTGTTGCGGCGCCACGGATCTTCAGTTCG 848

Qy   557 TGGATTGGTAAAAGTTTGTTTTTTGATGATAATATGCCATACAATCACCATCAAAT 616
Db   847 CCAACGGCCAGCAGCTTGTTCGAACTGGAACGCCCACGCGTTCCTTCCTCGATC 788

Qy   617 AAAATAAAATCTGCCGTTTGGTGGTCTAAGTAAGCTATATTCGTTGTGTGATTTT 676
Db   787 AGGGCCGAAAGC------------CTTGAAGTACCGGCGTTAACTCAGCAGCAGTG 750

Qy   677 GGGGCCAAGCGCTTGTCGTTGTCTTCGTCTAAATTGGCATCGTCTCTTGGTAG 736
Db   749 CGACCAGCACGGCACGGTATCGGCGGTCGGGTCA--GGGCATCGGCGCGCATTC 692

Qy   737 ATGATGGCGGTTTGGCAATTTTCAAAAAATACGCAATTTTTGTGCCAAATAATCCACCATG 796

RESULT 11
US-09-489-039A-5247/c
; Sequence 5247, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 5247
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-5247

Query Match   7.3%; Score 100.6; DB 3; Length 1416;
Best Local Similarity   51.1%; Pred. No. 4.1e-20;
Matches 337; Conservative   0; Mismatches 289; Indels 33; Gaps   3;

Qy   691 ACCACGACCTTGGCGCGGCGACCGCGGAAGATCCGGTCGTGCTTGGCCAGGTGGTAGTCAGCCATG 632
Db   :

Qy   722 GCATCGTGCTTGGTAGATGAGTGGCGGTGTTGGCAATTTCAAAAAATACGCAATTTTGTGCC 781
Db   797 CCATGGTGCCGATCAAGATGATCAGCGGAGCAGAGATTTAAAATTGTTGCCCCTTGAGCTGCC 856

Qy   782 AAATAATCCACCATGCC---ATCGTGCCGATCAAGATGATCAGCGGAGCAGATTTAAATT 838
Db   746 GCATCGTTCCGCGGTTCACCACGCGGGTTCTTCGGCCATTCTCGTAAATCCGACGCTTCGCGCG 687

Qy   839 GTTGCCCCTTGAGCTGCCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGC 898
Db   686 CGATACGTGCTGCGCAGCCCCAGGCGGGGTAGCGGTAGCGGTCCTCACGTTGACGGATG 627

Qy   899 ACCACCAAAATCCATATTTGGATCGTTAATAATTCAAGTGCAGGCGGTGCCAATATTACCG 958
Db   626 GTTGCCCCACCACGGCTGCGAGGTGGTTTCCAGCTGGAGGAGCTCCAGC 567

Qy   959 CCGACACCAACAATCATCGCCTGCACTGTTTCGCCATCTCGCCTAATGTGTAACACAGTG 1018
Db   566 ACATACAGCTCTCCGATCGGGTGCCAGCGTCCAGCGGAGGCGGTCTCAGGTAGTAACAGTG 510

Qy   1019 CTTTTGGCATTGAGCCTGTGATGCGGATGAGTTAGGTTGGTTGGTTGCTTGAGTTAGATCT 1078
Db   509 CCCACGCCGAGGTTAACTCAGCAGCAGTGGTCAAGCTCCAGGGTAGTAACAGTG 450

Qy   1079 CGTGCTTGAGTGTATCAAGTGAACAATTGAACATCGCTGATCACAGGAGAATGCCTTGAGCT 1138
Db   449 CTTTTACCGTTGAACCGGTAATGGCGATAATCGGCGCGCCCTGCCTCGGCG--------- 400

Qy   1037 GTGATGGGCGAGATTCGGTGTGTTTG 1061
Db   394 GTGATGGCGACGGATCGGGGCGCCTTCG 370

Qy   857 AAATTTGAGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCACCACAAAATCCATATTT 916
Db   857 AGGCCGATGCGGAGGTTTCCAGCTGGAAGCTGACAGCTCCAACAGCTCGATGTCG 512

Qy   917 TGGATCGTTAATAATTCAAGTGCAGGCGGTGCCAATATTACCGCCGACACCAACAATCATG 976
Db   571 AGGGCCATGCGAGGTTTCCAGCTGGAAGCTGACAGCTCCAACAGCTCGATGTCG 512

Qy   977 CCTGCACTGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGTCCGCGGGTGGTGCC 1036
Db   511 TCG---GCCAGCAGGTCGACGCGCCCGGGGTCCGGAGGTTGCCGCCACGCGTTG 455

Qy   1144 CGCAGGTCATGGAAAATCGGCGCGCCCCTTTGCCGCTCTCCGCGAGCAGCACCAGCTTGCCG 1085
Db   454 TCCGCGGGGCCACCGCCACCTTTCGCCCACACCAGGGTGGTCAACGGTGGTCTTCGCGGTTTGGAACCG 395

```
Db       399  ----------------GCAGAACAGTTGGATATGCCAAGGATCTCTAGCGGTCG   357
                              ||||||  || || |||| |||||||  |||| ||| |||
Qy      1139  TTGCGCGGGGACAATACTCGGCGTTCTTGGGTAAGTCGGCGGCTGATGAGTGATCTGATCG  1198
              |||  || ||  |   |    | ||| |||  |  ||| ||  |  |||
Db       356  GCCCGCTGCCCTGACGAAGGATGGGCGCCAGCGCAATCCCCGGACTGGCGACAATCAGGTCC   297
Qy      1199  GCTGAGCTTAAGAGTTCACTCGTGTCCAAAAATGACAATGAAGCCCTGTGGGCAGT  1258
              ||  |||  ||||   ||  ||   |||||   ||| |||   |||  ||    |
Db       296  GCCGCCAGCAGCGTATCCATTGAGTCCTCCGACGACACATGGCGCATTGGGTAAT   237
Qy      1259  TTATCTGCCAAGGTAGGATTGGGATTACCACCATCGATGAGATGCTTAACTTTGTGACCGTGATTG  1318
              |||| ||||  | || ||||||  | |||||| |  | ||| |||  ||||||||||  ||
Db       236  TTATCCAGCCCCGGAGGGGGTGTCCATCACGCGGCCCAGCCCAATTATGACGACTTTTTT   177
Qy      1319  ACCAAAAAATTCACAGCTGACAGTGCCAGAGCTGCCCAAACCGACAGCGGCCATATTTTTTT  1377
              ||||||   |||| |  |||| |  |  |||||||||||||   |   |||||||
Db       176  ATGAAAAAGTCCAGCGACATGACAGCGCGGTTGTCCATCACGCGGCCCAATTATGACGACTTTTTT   118
```

RESULT 12
US-09-790-988-1
; Sequence 1, Application US/09790988
; Patent No. 6632935
; GENERAL INFORMATION:
; APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATTORI, MASAHIRA
; APPLICANT: SAKAKI, YOSHIYUKI
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIOR APPLICATION NUMBER: JP2000-107160
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
;
; SEQ ID NO 1
; LENGTH: 640681
; TYPE: DNA
; ORGANISM: Buchnera sp.

Query Match   6.1%; Score 84.8; DB 3; Length 640681;
Best Local Similarity   50.2%; Pred. No. 8.4e-14;
Matches 289; Conservative   0; Mismatches 257; Indels 30; Gaps   2;

```
Qy       802  GTGCCGATCAAGATGATCAGCGGAGAGAGATTTAAAATTGTTGCCCCTGAGCTGCCAAATT   861
              ||  ||| ||  || |||    |  | |||  ||||||||||||   ||| ||| |||||
Db    240727  GTATCGATTGATGATGATCTTCGCTAATATTATTAAGAATAACTGCTATCTTAGATTTTAAATT   240786
Qy       862  TGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCACCACAAATCCATATTTTGGAT   921
              || ||  ||||| |||||| ||| || |   |||  |   ||  |||| |  |  |||
Db    240787  AAATGTATTTTCTAGTTGAAAACTAGACAGTTCTATTAT TATAAATCTGCTTCTT---T   240843
Qy       922  CGTTAATAATTCAAGTGCAGGCGTGCCAATAATTTACCGCGGACAACAATCATGCCTGC   981
              || |||||| ||||  | ||||  | |||| ||| ||| | ||| | ||| ||| |||
Db    240844  ATCAAGTATTTCTAGTACGGGGACACCTATATTATTACCACCTAGAAAAGCTTTTATATCCTGA   240903
Qy       982  ATGTTTTTGCCATCTCCGCCTACTATGTCGTGTAACAGTGTTTTTGGCATTGCTTGCTGCTGGT  1041
              |  |||||||||| |  |  |  | ||| ||  |  |||| |  |||||| ||  || | |
Db    240904  TTTTTTTGGCAATTTTTCAATCATGGGTAGTCAGTGCTACAGTTTTTTCAATTAGTGTATCAAATGA  240963
Qy      1042  GGGCCGATGAGTTGGTGTGTTCTGTGGCTTTAGATTCGTGCTTTGAGTGTATCAATGAA  1101
              ||  | || | |  | |  | | || |  | |  |  || |  |||| || |||||
Db    240964  TGAAATAATTAGGACAGGTCACTCTCTAG----------------AAAA   240996
Qy      1102  CAATTGAACATCCGCTGATCACACAGGAATGCCTTGGACCTTTGGGCCGCGACAATACTCGGGCCGT  1161
              ||||||  || |||||| |  || || ||| |  |||  ||| || |||| || | |  | |||
Db    240997  TAGTTCAATATCACTAATAATCTCAAAAACCTGCTTTAATTAAAATTAGGTTT   241056
Qy      1162  TCTTGGGTTAATGCCGGGGCTGAGTGATGATGATCTGGGCTGAGCTTAAGAGTTCACTGTC  1221
              |||  |    |  | ||||  |  | ||| |    || |  |  | ||||  || |  |  ||
```

RESULT 13
US-09-107-532A-1574/c
; Sequence 1574, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;   ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 893-5007
; TELEFAX: (781) 893-8277
; INFORMATION FOR SEQ ID NO: 1574:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1368 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1...1368
; LOCATION: (B) LOCATION 1...1368
; SEQUENCE DESCRIPTION: SEQ ID NO: 1574:
US-09-107-532A-1574

```
Db    241057  AAATGAAGAAGAAATACCCGGGACTTATAAACAATTAAGGTCTGATTCTGATGAGGTC   241116
Qy      1222  CAAGCTGCCAAAATGACAATGAAGGCCCTGTGGGCAGTTATCTGCCAAGGTAGGATTGGG  1281
Db    241117  TAAACTTCCTAAACTATATTCATTTGAGGAATTTAATAAAATTTGAAGGATGTTT   241176
Qy      1282  ATTACACATCGATGATGCTACTTTGTGACCGTGATTGACCAAAAAATTCACAGCTGACAG  1341
Db    241177  AGATTCATCAATAATTTTAGGTTTTTTATTCCTTTTTTTTAAAAAAAATTAAATGCAAGATAT  241236
Qy      1342  TCCAGAGCTGCCCAAACCGACGACGGGCCATATTTTTTT  1377
Db    241237  TCCTGTTAAACCATTCCTAAAAATTAATATATTTTTTT   241272
```

Query Match   6.0%; Score 83; DB 3; Length 1368;
Best Local Similarity   53.6%; Pred. No. 1e-14;
Matches 173; Conservative   0; Mismatches 150; Indels   0; Gaps   0;

```
Qy       312  ATGAATATGCCATAAACCGTGCCAAGCCCATCAAATTGCTGCAAATGGTTGAGCCGGATGTTTG   371
```

RESULT 14
US-09-252-991A-7787
Sequence 7787, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE:
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7787
; LENGTH: 567
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-7787
Query Match   5.1%; Score 69.8; DB 3; Length 567;
Best Local Similarity   55.0%; Pred. No. 6.9e-11;
Matches 137; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

```
Db 1069 ATAACACGACCTTAGACGTTTCAAATCCACTTAGTCTCGCTAAAATATTGG 1010
             |||                          ||||||
Qy  372 TACCTTTTGAATCATTAAAATATGCCTTGCTCTATCAATATCATCAATATATTCACGCGAT 431
Db 1009 TTGCTTTTGAATCATTATAAAAATTTTCTCGCCTTGGATTCTCCTACATATTGCGTACGAT 950
Qy  432 GCCGGTAGCCCTTTGAAAGTCTTAAGGTATCAACGATGCCATCCATGGGAAGATTGGCAG 491
Db  949 GTGGTACGCCCATGGAAAAAATGCAACGTTTCTCGTATACTTCATTGATATACCATACA 890
Qy  492 CTGTGCCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCCCTTTGATTTTA 551
Db  889 ATTTTGCTACAGAAAATAGGAGCTAGTGCATTTTCTACATTATTGCTTACCTGGTACGGCAA 830
Qy  552 ATTGATGATTGGTAAAAGTTTGTTTTTTGGATGATAATAATGCCATACCACCATCAT 611
Db  829 GTTCAGATATTCCATTATTTTTCTTTTTTATAGTAATTGAGTAATCGTCAGCGATACA 770
Qy  612 CAAATAAAATAAAAATCTGCCGTT 634
             ||| |||   |||||||
Db  769 CACCGTCTTCCAACACTTCTTTT 747
```

```
Qy  327 CCGTGCCCAAGCCCATCAATTGCTGCAATGGTTGAGCGATGTTTGTACCTTTTGAAATCAT 386
Db  262 CGGCACCCAGCCCCTTCGATCGCCCGGCACGTTGGTGCCCTTGAATCGT 321
Qy  387 TAAAATATGCCTTGCTCTATCAATATCATCAATATATTCACAGCGATGCCGGTAGCCCTTTGA 446
Db  322 CGTAGTAGTCTCACGCCCTCGGCGTACCCACTGGCGATGAGCCGCGGAAA 381
Qy  447 AAGTCTTAAGGGTATCAACGATGGCCATCCATGGGAAGATTGGCAGCTGCTGTGCCTAAGGCAA 506
Db  382 ACGCCTTCAGGCGGCCCGACGATGGCCATGGCCCAGCGCCA 441
Qy  507 GTGCAGATAAGGCATTGAGTAGGTTATGCTTGCCCTTTGATTTTAAATTGATTGGTTA 566
Db  442 GCCGCGGGCCGAGGCGCTGTTGGGAATAGTTGTGGCGCCCACGGATCTTCAGTTCGCCACGGGCA 501
Qy  567 AAAGTTTGT 575
Db  502 GCAGCTTGT 510
```

; Sequence 881, Application US/09221017B
; Patent No. 6444799
; GENERAL INFORMATION:
; APPLICANT: Ross, Bruce C.
; TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
; NUMBER OF SEQUENCES: 1120
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: MORRISON & FOERSTER
;   STREET: 755 PAGE MILL ROAD
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94304-1018
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: Windows
;   SOFTWARE: FastSEQ for Windows Version 2.0b
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/221,017B
;   FILING DATE: 23-DEC-1998
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PP1182
;   FILING DATE: 31-DEC-1997
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PP1546
;   FILING DATE: 30-JAN-1998
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PP2911
;   FILING DATE: 09-APR-1998
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/AU98/01023
;   FILING DATE: 10-DEC-1998
; ATTORNEY/AGENT INFORMATION:
;   NAME: Monroy, Gladys H
;   REGISTRATION NUMBER: 32,430
;   REFERENCE/DOCKET NUMBER: 27340-20021.00
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 650-813-5600
;   TELEFAX: 650-494-0792
;   TELEX: 706141
; INFORMATION FOR SEQ ID NO: 881:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 2402 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
;   ORGANISM: PORPHYROMONAS GINGIVALIS
; FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: 1...2402
US-09-221-017B-881

Query Match   4.9%; Score 67; DB 3; Length 2402;
Best Local Similarity   46.6%; Pred. No. 1.1e-09;
Matches 334; Conservative 0; Mismatches 365; Indels 18; Gaps 3;

```
Qy  355 GGTTGAGCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTCTATCAATATCATC 414
Db 2165 GGTAGAATTTACGTTCGTAGCTTTGGAGTCATTGATATAATCGACACCCTTTACGGCGAGC 2106
Qy  415 AATATATTCACAGCGATGCCGGTAGCCCTTTGAAAGTCTTAAGGTATCAACGATGCCATC 474
Db 2105 AAICTTTTCCCAGTCGATGCGGGTACATTCTTGAAAAATCCTGCAAGCCAAGGCTTCTCGAAGGCTC 2046
Qy  475 CATGGGAAGATTGGCAGCTGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATG 534
Db 2045 ATTCTTAATATCCATAGCTTTGGGCAGCAATGGCAGTTGCCATAGCAATTGTGCCTGTGTG 1986
```

```
Qy    535 CTTGCCTTTGATTTTTAATTGATGGATTGGTAAAGTTTGTTTTTTGATGATATAATGC 594
          |||||
Db   1985 CATGCCGGACAGAGCTAAAAGTTCCTCATCCATAACAAAAGGTGAAATTAAGATTCATGAC 1926

Qy    595 CATACCAATCACCATCATCAAATAAATAAAAATCTGCCGTTTGGTGGCTAAGTAAGCTATT 654
          |||        |||
Db   1925 AACCAATTCGTTCTTTTCATTGATCCAAGCAGTTGTATTATCCGTTCGAGCTTCCATTGC 1866

Qy    655 TGTTGTGATAATGGTGTGATTTGGGGCAACGCTTGTCTCAGTGGTGTCAAGCATTGCTTGGC 714
          |||
Db   1865 GAACGGGTAGAAGAC--------GGGCTACAGGAGGATGTTCGGCGCTACCCATCGGCCTGAT 1815

Qy    715 TAAATTGGCCATCGTCCTTCTTGGTAGATGGATG----------GCGGTTTGGCAATTTTCAAAAATACG 768
          ||         |||||||||
Db   1814 AAAAAGGATCATCCTCCCAGTAGATGAAGCAAATCCTCGGGGTTGCTGATTCCGAGTGGATTCT 1755

Qy    769 CAATTT----TTGTGCCAAATAATCCACCATGCCGATCAAGATGATCAGCGGGA 825
          ||||       |||
Db   1754 CATTTTCGCTTCTGCATAAAGCTCGAATCGGTGATCGTATCGGATCTAAAATGATCCGGGTGT 1695

Qy    826 GAGATTTAAAAATTGTTGCCCTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACT 885
          |||             |||
Db   1694 AATATTGAGCAGAATGGCCACATTAGCCCTGAAGTCGTACATGTTGTCCAGTTGAAAACT 1635

Qy    886 TGACAGCTTCAAGCACCACCAAAATCCATATTTTGGATCGTTAATAAATTCAAGTGCAGGCGT 945
          |||
Db   1634 GCTCAACTCTATTACATAATAATAAGGATGAGGATCGTAAGCGACCTGTCTGGCCAGACTAAA 1575

Qy    946 GCCAATATTACCGCCGACACCAACAAATCATGCCTGCATGTTTTCCATCTCGCCTACTAA 1005
          ||     |||
Db   1574 ACCGACAATTGCCTGCCAAAACCTACATCCAAAACCGGGCCTTACAGAGGCGTGTGATACAACCA 1515

Qy   1006 TGTCCGTAACAGTGCTTTTTGGGCAATTTGAGCCTGTGATGGCCGATGAATTGGTGTGTTTTGT 1062
          |||
Db   1514 CATGGTTGTTGTTGTCTTGCCATTACTACCCGTGATGCACACCATAAAATGCATCGGT 1458
```

THIS PAGE BLANK (USPTO)

OM nucleic - nucleic search, using sw model

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

Run on: May 18, 2006, 09:31:41 ; Search time 1686 Seconds
                    (without alignments)
                    10057.495 Million cell updates/sec

Title: US-10-672-787-35_COPY_11357_12736
Perfect score: 1380
Sequence: 1 taaagcgtgaaccatgttaa..........acgaccggcatattctttga 1380

Scoring table: IDENTITY_NUC
                    Gapop 10.0 , Gapext 1.0
Searched: 18892170 seqs, 614817638 residues

Total number of hits satisfying chosen parameters: 27508734

Minimum DB seq length: 22
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications NA Main:*
                    1: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US07_PUBCOMB.seq:*
                    2: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US08_PUBCOMB.seq:*
                    3: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US09A_PUBCOMB.seq:*
                    4: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US09B_PUBCOMB.seq:*
                    5: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US09C_PUBCOMB.seq:*
                    6: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10A_PUBCOMB.seq:*
                    7: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10B_PUBCOMB.seq:*
                    8: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10C_PUBCOMB.seq:*
                    9: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10D_PUBCOMB.seq:*
                    10: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10E_PUBCOMB.seq:*
                    11: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10F_PUBCOMB.seq:*
                    12: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US10G_PUBCOMB.seq:*
                    13: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US11A_PUBCOMB.seq:*
                    14: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US11B_PUBCOMB.seq:*
                    15: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US11C_PUBCOMB.seq:*
                    16: /EMC_Celerra_SIDS3/ptcodata/2/pubpna/US11D_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	1380	100.0	1398	8	US-10-282-122A-27195	Sequence 27195, A
c 2	1380	100.0	96109	8	US-10-677-787-35	Sequence 35, Appl
c 3	231	16.7	1344	8	US-10-282-122A-8940	Sequence 8940, Ap
c 4	160.4	11.6	1314	8	US-09-815-242-7090	Sequence 7090, Ap
c 5	160.4	11.6	1314	10	US-10-282-122A-22236	Sequence 22236, A
c 6	160.4	11.6	2925	15	US-11-194-246-220	Sequence 220, App
c 7	160.4	11.6	11498	4	US-09-754-468-44	Sequence 44, Appl
c 8	160.4	11.6	1830121	8	US-09-754-468-44	Sequence 44, Appl
c 9	160.4	11.6	1830121	8	US-10-329-670-1	Sequence 1, Appl
c 10	160.4	11.6	1830121	9	US-10-158-865-1	Sequence 1, Appl
c 11	160.4	11.2	14324	10	US-10-981-687-1	Sequence 1, Appli
c 12	154.4	11.6	908766	10	US-10-795-159-569	Sequence 569, App
c 13	154.4	11.1	1281	8	US-10-795-159-685	Sequence 685, App
c 14	152.8	11.1	1314	10	US-10-282-122A-25165	Sequence 25165, A
c 15	152.4	11.1	1305	10	US-10-958-216-402	Sequence 402, App
c 16	142.2	10.3	1305	8	US-10-282-122A-30713	Sequence 30713, A
c 17	139.6	10.1	1317	3	US-09-741-669-259	Sequence 259, App
c 18	139.6	10.1	1317	3	US-09-815-242-5932	Sequence 5932, Ap
c 19	139.6	10.1	1317	3	US-10-282-122A-20229	Sequence 20229, A
c 20	139.6	10.1	28277	4	US-09-754-468-2	Sequence 2, Appli
c 21	132.8	9.6	2850	15	US-11-194-246-205	Sequence 205, App
c 22	132.8	9.6	5640	15	US-11-194-246-205	Sequence 205, App
c 23	123.6	9.0	1308	8	US-10-282-122A-32629	Sequence 32629, A
c 24	117.8	8.5	1182	8	US-10-282-122A-33528	Sequence 33528, A
c 25	117	8.5	1347	8	US-09-815-242-7912	Sequence 7912, Ap
c 26	117	8.5	1347	8	US-10-282-122A-30468	Sequence 30468, A
c 27	113	8.2	1404	8	US-10-282-122A-41274	Sequence 41274, A
c 28	111	8.1	1317	8	US-10-282-122A-39003	Sequence 39003, A
c 29	110.8	8.0	4942	10	US-10-450-763-25620	Sequence 25620, A
c 30	110.8	8.0	4944	10	US-10-450-763-25077	Sequence 25077, A
c 31	110.8	8.0	4944	10	US-10-450-763-25985	Sequence 25985, A
c 32	110	8.0	1350	8	US-10-282-122A-31716	Sequence 31716, A
c 33	109.4	7.9	1317	8	US-09-815-242-9978	Sequence 9978, Ap
c 34	109.4	7.9	1317	8	US-10-282-122A-39835	Sequence 39835, A
c 35	103.2	7.5	1314	8	US-10-282-122A-23400	Sequence 23400, A
c 36	94.2	6.8	2058	10	US-10-282-122A-19656	Sequence 19656, A
c 37	92.8	6.7	4831	10	US-10-450-763-8816	Sequence 8816, Ap
c 38	92.8	6.7	4831	10	US-10-450-763-30346	Sequence 30346, A
c 39	92.6	6.7	1317	10	US-10-282-122A-41740	Sequence 41740, A
c 40	89.6	6.5	3117	10	US-10-450-763-4899	Sequence 4899, Ap
c 41	89.6	6.5	3117	10	US-10-450-763-25044	Sequence 25044, A
c 42	89.6	6.5	3117	10	US-11-282-122A-17455	Sequence 17455, A
c 43	84.8	6.4	1353	3	US-09-790-988-1	Sequence 1, Appli
c 44	84.8	6.0	1359	8	US-10-282-122A-21621	Sequence 21621, A
c 45	81	5.9	1071	8	US-10-282-122A-37050	Sequence 37050, A

ALIGNMENTS

RESULT 1
US-10-282-122A-27195/c
Sequence 27195, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:

APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308

```
;  PRIOR FILING DATE: 2001-02-16
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 78614
;  SOFTWARE: Patentin version 3.1
;  SEQ ID NO 27195
;  LENGTH: 1398
;  TYPE: DNA
;  ORGANISM: Moraxella catarrhalis
US-10-282-122A-27195

Query Match          100.0%;  Score 1380;  DB 8;  Length 1398;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1380;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATAATAATTTTTAAATTGATC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1383 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATAATAATTTTTAAATTGATC 1324

Qy      61 CATGCTAGCACAAGCAGGTGACAAAAGACACAGCCCTAGGGGTGCTTGCTTGACTGGCTGCCAA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1323 CATGCTAGCACAAGCAGGTGACAAAAGACACAGCCCTAGGGGTGCTTGCTTGACTGGCTGCCAA 1264

Qy     121 TACTGCCCGGATCAAGTGTACCAACCTGATATATATTTTAGATGCCAAATGTGCATCACCTTT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1263 TACTGCCCGGATCAAGTGTACCAACCTGATATATATTTTAGATGCCAAATGTGCATCACCTTT 1204

Qy     181 GAGTAAAATCTTGCTCAATTATCGGGGCATCTTGACCGGATTAAATACACTACTATCCACCCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1203 GAGTAAAATCTTGCTCAATTATCGGGGCATCTTGACCGGATTAAATACACTACTATCCACCCA 1144

Qy     241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAAT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1143 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAAT 1084

Qy     301 CAATGCCAGTGAATGATGATATATGCCATAAACCCTGCCAAGCCCATCAAATTGCTGCAATGGTTGA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1083 CAATGCCAGTGAATGATGATATATGCCATAAACCCTGCCAAGCCCATCAAATTGCTGCAATGGTTGA 1024

Qy     361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGGCTCATATCAAATATCATCAATATA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1023 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGGCTCATATCAAATATCATCAATATA 964

Qy     421 TTCACAGCGGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCGCATGGCATCCATGGG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     963 TTCACAGCGGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCGCATGGCATCCATGGG 904

Qy     481 AAGATTGGCAGCTGTGCCTAAGGCAAGTCCAGATAAGGCAGTCCCTTTGAAAGTAGGTTGCC 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 AAGATTGGCAGCTGTGCCTAAGGCAAGTCCAGATAAGGCAGTCCCTTTGAAAGTAGGTTGCC 844

Qy     541 TTTGATTTTTTAATTGATGATGGATTGGTGATAAAAGTTGTTTTTTTGATGATAATGCCTATACC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 TTTGATTTTTTAATTGATGATGGATTGGTGATAAAAGTTGTTTTTTTGATGATAATGCCTATACC 784

Qy     601 ATCACCATCATCAAATAAATAAAAATCTGCCGTTGGGTTTGGGTGGCTAAGTAAGCTATTTGTTGT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 ATCACCATCATCAAATAAATAAAAATCTGCCGTTGGGTTTGGGTGGCTAAGTAAGCTATTTGTTGT 724

Qy     661 GATAATGGTTGATTGGGGGCAACGGCTTGTTCAGTGTGTTCAAGCATTGCTTGGCTAAATT 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 GATAATGGTTGATTGGGGGCAACGGCTTGTTCAGTGTGTTCAAGCATTGCTTGGCTAAATT 664

Qy     721 GGCCATCGTCTTGGTAGATAATGGGGCGGTTTGGCAAATTTCAAAAAAATACCGCAAATTTGTGTC 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 GGCCATCGTCTTGGTAGATAATGGGGCGGTTTGGCAAATTTCAAAAAAATACCGCAAATTTGTGTC 604

Qy     781 CAAATAATCCACCATGCCATGCCGGATCGGATCAAGATTAAAGATTTAAAAATTGT 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     603 CAAATAATCCACCATGCCATGCCGGATCGGATCAAGATTAAAGATTTAAAAATTGT 544

Qy     841 TGCCCCTTGAGCTGCCAAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCAC 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 TGCCCCTTGAGCTGCCAAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCAC 484
```

```
RESULT 2
US-10-672-787-35
;  Sequence 35, Application No. US20040067554A1
;  Publication No. US20040067554A1
;  GENERAL INFORMATION:
;  APPLICANT: LAGACE, Robert, E.
;  APPLICANT: PATTERSON, Chandra
;  APPLICANT: BERG, Kim, L.
;  TITLE OF INVENTION: NUCLEOTIDE SEQUENCES OF MORAXELLA CATARRHALIS GENOME
;  FILE REFERENCE: ELITRA.025C1
;  CURRENT APPLICATION NUMBER: US/10/672,787
;  CURRENT FILING DATE: 2003-09-26
;  PRIOR APPLICATION NUMBER: 09/596,002
;  PRIOR FILING DATE: 2000-06-16
;  NUMBER OF SEQ ID NOS: 41
;  SOFTWARE: PERL Program
;  SEQ ID NO 35
;  LENGTH: 96109
;  TYPE: DNA
;  ORGANISM: Moraxella catarrhalis
US-10-672-787-35

Query Match          100.0%;  Score 1380;  DB 8;  Length 96109;
Best Local Similarity   100.0%;  Pred. No. 0;
Matches 1380;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy       1 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATAATAATTTTTAAATTGATC 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11357 TAAAGCGTGAACCATGTTAACAAAACTTATCACCTCGCTCATTATAATAATTTTTAAATTGATC 11416

Qy      61 CATGCTAGCACAAGCAGGTGACAAAAGACACAGCCCTAGGGGTGCTTGCTTGACTGGCTGCCAA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11417 CATGCTAGCACAAGCAGGTGACAAAAGACACAGCCCTAGGGGTGCTTGCTTGACTGGCTGCCAA 11476

Qy     121 TACTGCCCGGATCAAGTGTACCAACCTGATATATATTTTAGATGCCAAATGTGCATCACCTTT 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   11477 TACTGCCCGGATCAAGTGTACCAACCTGATATATATTTTAGATGCCAAATGTGCATCACCTTT 11536

Qy     181 GAGTAAAATCTTGCTCAATTATCGGGGCATCTTGACCGGATTAAATACACTACTATCCACCCA 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 GGCGCGGCGACAATACTCCGGCGCGTTCTTGGGTTAATGCCGGGGGCTGATGATGAGTCGGC 4

Qy     901 CACCAAAATCCATATTTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATATTACCGCC 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 CACCAAAATCCATATTTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATATTACCGCC 424

Qy     961 GACACCAACAAATCATGCCTGCACAGTGTTTGCCATCTCGCCTACTAAGTCGTCGTAACAGTGCT 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GACACCAACAAATCATGCCTGCACAGTGTTTGCCATCTCGCCTACTAAGTCGTCGTAACAGTGCT 364

Qy    1021 TTTGCCATTTGAGCCTGTGAGTGACCGATGATTGGGCCGGATGATTGGGTTGTTGTTCTTAGATCTCG 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 TTTGCCATTTGAGCCTGTGAGTGACCGATGATTGGGCCGGATGATTGGGTTGTTGTTCTTAGATCTCG 304

Qy    1081 TGCTTTGAGTGTATCAAATGAACAAATTGAACATCGCTGATCACAGCAATGCCTTGAGCTTT 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 TGCTTTGAGTGTATCAAATGAACAAATTGAACATCGCTGATCACAGCAATGCCTTGAGCTTT 244

Qy    1141 GGCGCGGCGACAATACTCCGGCGCGTTCTTGGGTTAATGCCGGGGGCTGATGATGAGTCGGC 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 GGCGCGGCGACAATACTCCGGCGCGTTCTTGGGTTAATGCCGGGGGCTGATGATGAGTCGGC 184

Qy    1201 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAAATGAACAGCGCCCTGTGGGCGCAGTTT 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAAATGAACAGCGCCCTGTGGGCGCAGTTT 124

Qy    1261 ATCTCGCCAAGGTAGGATTGGGATTACCATGGCATGCTAACTTTGTGACCGTGATTGAC 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 ATCTCGCCAAGGTAGGATTGGGATTACCATGGCATGCTAACTTTGTGACCGTGATTGAC 64

Qy    1321 CAAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAAACCGACGACGGGCCATATTTTTTTGA 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 CAAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAAACCGACGACGGGCCATATTTTTTTGA 4

Qy     181 GAGTAAAATCTTGCTCAATTATCGGGGCATCTTGACCGGATTAAATACACTACTATCCACCCA 240
```

```
Db  11537  GAGTAAATCTTGCTCAAATTATCGGGGGCATCTTGACCGGATTAAAATACACTATCCACCCA  11596
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy    241  CTGCATAACATATCGCAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT       300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11597  CTGCATAACATATCGCAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAT      11656

Qy    301  CAATGCCAGTGATGATGATGATGCCATAAAACGTGCCAAGCCCAATCAATTGCTGCAATGGTTGA  360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11657  CAATGCCAGTGATGATGATGATGCCATAAAACGTGCCAAGCCCATCAATTGCTGCAATGGTTGA  11716

Qy    361  GCCGGATGTTTGTACCTTTTGAAATCATTAAAATATCCCTTGCTATCAATATCAATCAATATA  420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11717  GCCGGATGTTTGTACCTTTTGAAATCATTAAAAAATATGCCTTGCTATCAATATCAATCAATATA  11776

Qy    421  TTCACAGCGGATGCGGTAGCCCCTTTTGAAAGTCTTAAGCATCTTAAGGGTATCAAGCATTCATCAATGG  480
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11777  TTCACAGCGATGCGGTAGCCCCTTTTGAAAGTCTTAAGGGTATCAACCATGGCATCCATGGG  11836

Qy    481  AAGATTGGCAGCTGTGCCTAAGGCAAGTGCCAGATAAGGCCATTGAGTAGTTATGCTTTGCC  540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11837  AAGATTGGCAGCGCTGTGCCTAAGGCAAGTGCGAGATAAGGCCATTGAGTAGGTTATGCTTTGCC  11896

Qy    541  TTTGAATTTTTAATTGATGATGGATTGGTAAAAGTTGTTTTTTTTGTTTTTTGATGATATAATGCCATACC  600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11897  TTTGAATTTTTAATTGATGGATTGGTAAAAGTTGTTTTTTTGATGATATAAATGCCATACC  11956

Qy    601  ATCACCATCATCAAATAAATAAATATCTGCCCGTTGTGGTGCTAAGTAACGTATTTGTTGT  660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  11957  ATCACCATCATCAAATAAATAAATATGCCTTGCTATCAAGTTAAGCTATTTGTTTGTTGT  12016

Qy    661  GATAAAGGTCTGATTGGGGGCAACCGCTTGTCTCGGTGGTTGGTGGTCTCAAGCGTTGGGCTAAATT  720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12017  GATAATGGTGTGATTGGGGGCAACGCTTGTCTCAGTGTGGTCTCAAGCATTGCTTGGCTAAATT  12076

Qy    721  GGCATCGTCTTGGTAGATGATTGGCCGTTTGGCAATTTCAAAAAAATACGCAAATTTTTGTGC  780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12077  GGCATCGTCTTGGTAGATGAGTGGCGGTTTGGCAATTTCAAAAAATACGCAAATTTTGTGC  12136

Qy    781  CAAATAATCCACACATGCCATCGTCGCCGATCAAGATGATCAGCGGAGAGATTTAAATTGT  840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12137  CAAATAATCCACATGCCATCGTCGCCGATCAAGATGATCAGCGGAGAGATTTAAATTGT  12196

Qy    841  TGCCCCTTGAGCTGCCAAAATTTGAGATATGCTCAAGCTGGTGGTTCAGTGTGACAGCTCAAGCAC  900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12197  TGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGGTGGTTCAGTTGACAGCTCAAGCAC  12256

Qy    901  CACCAAATCCATATTTTGGATTCGTTAATAATTCAAGTGCAGGCGGTGCCAATATTACCGCC  960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12257  CACCAAATCCATATTTTGGATTCGTTAATAATTCAAGTGCAGGCGGTGCCAATATTACCGCC  12316

Qy    961  GACACCAACAAATACATGCCTCGACATGTTTTCGCCATCTTACTAATGTCGTAACAGTGCT  1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12317  GACACCAACAAATCAATGCCCTCGACATGTTTTCGCCATCTTGCTACTAATGTCGTAACAGTGCT  12376

Qy   1021  TTTGGCATTTGAGCCTGTGAATGGGCCGATGATTGGTGGTAGTGGTTGCTTGAGTTAGATCTCG  1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12377  TTTGGCATTTGAGCCTGTGATGGGCGGATGATTGGTGGTGTTGTTGCTTGAGTTAGATCTCG  12436

Qy   1081  TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTGCTTGAGCTTT  1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12437  TGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTGCTTGAGCTTT  12496

Qy   1141  GGGGGCGGACAATACTTCCGGCGGTTCTTGGGTTAATATCCCGGCGCTGATGATGATCTCGGC  1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12497  GGGGGCGGACAATACTTCCGGCGGTTCTTGGGTTAATGCCGGCGGCTGATGATGATCGGGC  12556

Qy   1201  TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAAATGACAATGAACGGCCTGTGGGGCAGTTT  1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12557  TGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGGCCTGTGGGGCAGTTT  12616

Qy   1261  ATCTGCCAAGGTAGGATTGGGGATTACCATCGATGATGATGCCATAAACGACGACGGCGCATATTTTGTGACCGTGATTGAC  1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12617  ATCTGCCAAGGTAGGATTGGGGATTACCATCGATGATGATGCCATAAACGACGACGGCGCATATTTTGTGACCGTGATTGAC  12676

Qy   1321  CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAGAGCTGCCCAAAACCGACGACGGCGCATATTTTTTGA  1380
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  12677  CAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAGAGCTGCCCAAAACCGACGACGGCGCATATTTTTTGA  12736
```

RESULT 3

US-10-282-122A-8940/c
; Sequence 8940, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-02-20
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; Remaining Prior Application data removed - See File Wrapper or PALM.
;
; SEQ ID NO 8940
; LENGTH: 1344
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii

US-10-282-122A-8940

Query Match   16.7%; Score 231; DB 8; Length 1344;
Best Local Similarity   52.4%; Pred. No. 4.1e-53;
Matches   721; Conservative   0; Mismatches 590; Indels   66; Gaps   7;

```
Qy      2  AAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATTAATTTAAAATTGATCC    61
           ||||||| || ||||    ||| || | |  | ||   | |||| | | || ||| |||||
Db   1340  AACGAATTGACGCACGCAACAAAACTGCTGACCGGTCATTATTAATACTTTAAAACATATCA  1281

Qy     62  ATGCTAGCACAAGCAGTGACAAAAGCACACAGCCCTAGCGGTGCTTTGACTGGCTGCCAAT   121
           |||||| ||||| || |||||||| |   | |  ||  |  | |   |||  |  |||||
Db   1280  AAACTTGCACATGCACGCTGGTGAT------AGCAATACCACATCTTCAGCTTGTTCAC   1229

Qy    122  ACTGCCGCATACAAAGTGTACCAAACCTGATATATTTTTAGATGCCAAAATGTGCATCACCTTTG   181
           ||||| | || ||||| | ||  | |  || |||||| |||| || |||| ||||| | ||  |
Db   1228  GTTGACACAGCTCGACGAGCTTCTTCTTTAAGCGGTGCTGCCATGTAAAAAATTTTAGTTGCGCCT  1170
```

```
Qy     182 AGTAAAATCTTGTGCTCAAATTATCGGGGGCATCTTTGACCGATTAAAATACACTATCCACCCAC 241
           |||| |||| ||| || |  |                          |||  ||   || |
Db    1169 TGAATGGCTTGCTCGGATGACCGGCCGCCATCTTCACCAATCAACTACCAACTTTGGCATAT 1110

Qy     242 TGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTACCAAGCCCACCCAAAATC 301
           |||  |  |  |  ||||   | ||                    ||| ||
Db    1109 TTTTCGATAGAGAACGTAAAGGAGAAAAATCTTGCCCTTGCCTGTCCACTTAAAAATA 1050

Qy     302 AATGCCA------GTGATGATATGGCATAAAACGTGCCAAGCCCATCAATTGCTGCAATG 355
           | |||||      ||| |  | |||  |     ||||||||| ||||||||||||||
Db    1049 AGTGCAACCTTACCTTTTTTCACTTCAATGGCAGCGACCTAAGCCATCAATTGCTGCAAGT 990

Qy     356 GTTGAGCCGATGTTTGTACCTTTTTTGAATCATTAAAAATATGCCTTGCTATCAATATCATCA 415
           | ||| ||| ||||| |||||||    ||  ||   ||||   |||  |   |  ||   |||
Db     989 GTAGCACAACATTGGTACCTTTAGAGTCATTATATAAATACGGACACCACGAACAGTTTTA 930

Qy     416 ATATATTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCC 475
           |||| | |||| || || | |   | ||  ||| || ||| | |  ||||| || ||||
Db     929 ACGTACTCACAGCGATGCTTAATCCTTTAAAATCGTTTAAAGTTTCAAGGCATTGACTCC 870

Qy     476 ATGGGAAGAGATTGGGCAGCTGGTCGTGCCTAAGGCAAGTGGCAAGAAGGCAGTAAGGATTATGC 535
           |  ||| | || |   ||| |   || || |||   | || ||| | |    ||| |||
Db     869 ATAGGTAAACCAATGCCTCCCCCCAATGCTAAAACACGCTAAAGCATTGGGCTACGTTGTGC 810

Qy     536 TTGCCTTTGATTTTTTAATTGATGATGGATTGGTGTAAACAGTTTGTTTTTTTGATGATATAAATGCC 595
           ||||  |||| | ||| | |||| |  ||  ||| | |  || || ||||   ||| | |||  |
Db     809 ATACCTTGAATATATAAAATCCGAGCTTTTAATTAACCACGACAAGCCAAGTTTTTTTTTT 750

Qy     596 ATACCATCACCATCATCAAATAAAATAAAATCTCGCCGTTTGGGTGGCTAAGTAAGCTATTT 655
           ||| | ||| || ||  ||||     | |  || || | |||  ||| |||
Db     749 AGCGTACCGTCGGGCATCTCTTTAAAACGCCATAATTGATT------TAAATCTGTGCGT 698

Qy     656 GTTGTGATAAGTGTGATTGGGGGCAACGCTTGTCAGTGTGGTCAAGCAATTGCTTGGGCT 715
           | ||  | || | || ||   ||   |  || |||| |   | ||||| || | ||  |
Db     697 TTAAACCAAAGCTTTGCATTGGTGTAATCGTGGAACAAGTGGGCTACGTTTGTC------- 647

Qy     716 AAATTGGCATCGTCTTGGGTAGATGATGGGCGGTTCGCAATTTCAAAAAATACGCAAATTTT 775
           |||| |   || || ||| |||| |  |||  ||| | ||  | |||   || |  |||
Db     646 ---AAGCATCGTCTCGGTTAAATACAACTTTTTTAGCGCCTTGAAATACCACGATGTTTTT 591

Qy     776 TGTGCCAAATAATCCACCACCATCCGTCGCCGATCAGCGGAGAGATTTAAA 835
           || |   |||  || || | |  | | | ||||   | || || ||||  |
Db     590 GCTGGTGATACCCGCAGCATATTTCCATGACGGTCTAAAATGGTCTTCATCATTATTGAGA 531

Qy     836 ATTGTTGCCCCTTGAGCTGCCAAATTTGGATAATGTCTCAAGCTGAAAAACTTGACAGCTCA 895
           | | ||| || | |  |  ||| || |||| ||| | ||||| ||| |||  || |||
Db     530 ACCACTGCTTACCTCAGCGTTTAAGTGAGAGGTGGTTTCTAACTGAAGAGCTTGATTAACTCA 471

Qy     896 AGCACCACCAAAATCCATATTTTTGGATCGTTAATAAATTCAAGTGCAAGGCTGCCAATATTA 955
           ||||| |  ||||| | ||| ||| |  ||| || |    |  ||   ||  || |||  |
Db     470 AGGACTAGTAGCTC---TGGTTGATCTTAAGTAAAATCTAAACGCTGGTCGACCAAGATTG 414

Qy     956 CCGGCCGACACCAACAAATCATCGCTGCACTGTTTTGCCATCTCGCCTTACTAATTGTCGTAACA 1015
           |||||||  ||| || || | | ||||| |  | ||| ||| |  |    | ||| || |  ||
Db     413 CCGGCTACTGCAACACTTTTCTTACCTGCAATTTCTTACCTGAAAAGTGGTTACA 354

Qy    1016 GTGCTTTTTGGCATTTGAGCCTGTGATGGGATGAGTGTTTGTTGCTGAGTTAGA 1075
           ||| || ||| || | ||  ||| ||||||| |||| | ||| ||| | ||||
Db     353 GTACTTTTTGCATTTGGAACCTGTAATCAACAATCGGCCACATCAGTCGGC---- 304

Qy    1076 TCTCGTGCTTTGAGTGTATCAAATGAACAATGCAGGAATGGCCTTGA 1135
           || || |||| || | |||||||| || |||| || ||||| |||||
Db     303 ------AGGACGCAAATATTGGATTGAACAATGCGGTGATATACCTTTA 261

Qy    1136 GCTTTGGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCCGCTGATGATGATTCTGA 1195
           | ||| |||| ||  ||   || |   |||||| ||||| | |    ||| ||| ||| |
Db     260 GCAATAGCTGCCGAATTTCCGGTAATGTGTGCAAGGCGTTAAAATAATCTCT 201

Qy    1196 TCCGGCTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGAACACGCCTGTGGGC 1255
           || |||| || | ||   | |||  | ||||| |  ||||| || | || | ||  |
Db     200 TCTCCTTGTTAATAAATTCTTGACCAAGTGACCAAAAACTGGTTTTTAACGCCAGCCGGA 141

Qy    1256 AGTTTATCTGCCAAGGTAGGAATTGGGGGATTACCATCGATGATGTGCTAACTTTGTGACCGTGA 1315
           | ||||| ||||| | | || |||| ||  || | ||| ||| ||| | |||  ||| |||
Db     140 ATCTGATTCGGTGTCCGGGAGGTGTGGGGGGAGGTCCGTTACAGACGAACTTGGTTAGCCTTGT 81

Qy    1316 TTGACCAAAAAAATTCACAGCTGTCGACAGTGGGGGGGGGGAGTGGGGGGAGGTCCGTTACAGACGAGGGGCATAT 1372
           | ||||| ||||||||||| || |||  |||||  | |  |||| || |   || ||| || |   || |
Db      80 TCATGCGAGGAAAATTTACAGCAGGAAACACCTGATATATCCCAAGCCTGCTACAAACTTTT 24

Qy       1 TAAAGGGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 60
           |||   |||||  ||||||| | |  || |  |  | | |  |||||   | ||||||||
Db    1308 TAAAACATTGAGCTAAATGCGTAAAATTCTTGCCGCGGCCTTTTCAAAAGAGAAAACTGATC 1249

Qy      61 CATGCTAGCACAAGCAAGGTGACAAAAGCACAGCCCTAGCGGTGCTGCCA 120
           |||||||||| |||  | |||||| ||||| ||||| ||| ||||||||||
Db    1248 GAGACTTGCACAAGCAGGCGACAATAATACCA------TATCTCGCTTTGACTTGTGTCCAA 1201

Qy     121 TACTGCCCGATCAAGTGTACCAACCTGATATATATTTTAGATGCCAAATGTGCCATCACCTTT 180
           ||||||| || ||||||||||||||| ||||| || | |||||| | |||| ||| |||
Db    1200 TGTTGGGGCGTAAAAATTCTATCGCTTGTTCCATTGTATCGAACATA------ATAACTTTG 1147

Qy     181 GAGTGAAATCTTGCTCAATTATCGGGGGCATCTTGACCGGATTAAATACACAATCATCCACCA 240
           ||| ||||||||| || |||| | | | || |||||||||| ||||| || | | |  |||
Db    1146 CGATGAAAATTTTGCAAGCAGCGCGCACCCATCTCGACCAAAACAATAACAAATAATGTGTGG 1087

Qy     241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTACCAAGCCTTACCACCCAAAAT 300
           |||   || |  ||||   || ||||   |  | ||||| | | |  |    |||   ||||
Db     241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTACCAAGCCCACCCAAAAT 300
```

Query Match   11.6%; Score 160.4; DB 3; Length 1314;
Best Local Similarity 49.6%; Pred. No. 2.1e-33;
Matches 683; Conservative   0; Mismatches 606; Indels 89; Gaps 7;

RESULT 5
US-10-282-1222A-22236/c
Sequence 22236, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
;
; SEQ ID NO 22236
; LENGTH: 1314
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
US-10-282-1222A-22236

Query Match   11.6%; Score 160.4; DB 8; Length 1314;
Best Local Similarity   49.6%; Pred. No. 2.1e-33;
Matches 683; Conservative   0; Mismatches 606; Indels   89; Gaps   7;

Qy   1 TAAAGCGTGAACCATCGTTAACAAACTTATCACCTCGCTCATTATAATTTTAAATTGATC 60
            ||| || ||| | |||| ||||||| || | || || |||||| | | ||||||||
Db   1308 TAAACATTGAGCTAAATGCGTAAATTCTTCGCCGCGCCTTTTCAAAAGAAGCAAACTGATC 1249

Qy   61 CATGCTAGCACAAGCACAGGTGACAAAAGCACAGCCCTAGCGCGTGCTTTGACTGCTCCAA 120
            |||| | |||||||||||| |||| | | || ||||| ||
Db   1248 GAGACTTGCACAAGCACAGGCGACAATAATACCA------------TATCTCCGCTTTGCAA 1201

Qy   121 TACTGCCCGATCAAGGTGTACCAACCTGAATATATTTTAGATCCCAAATGTGCATCACCTTTT 180
            | || ||||| | | | | | |||| ||||||
Db   1200 TGTTGGGCGTAAAAATTCTATCGCTTGTTCATTGTATCGAAAGTT 1147

Qy   181 GAGTAAAATCTTGCTCAATTATCGGGGGCCATCTTGACCGATTAAATACACACTATTCCACCA 240
            |||||| |||| || || | || || ||||| || ||| ||||||||||| || ||
Db   1146 CGATGAAAATTTTGCAAGCACGCGCACCATCTCGACCAAAACAACAATAACATTAAGTGTGG 1087

Qy   241 CTGCATAAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTTACCAAGCCACCCAACCAAAAT 300
            |||||||| ||||||||||| || || |||||||||| ||| | || ||||| || |||||
Db   76 AGAGAATAATCCACACAAGAAAGACCTGTTTTGCAAGCCGATGATTAGTAATATATTTTTT 19

Db   1086 CAATGCCAGTGATGAATGCTAATTCTTCAGCTTAATTCTCTGAAAAATCAGCCCCTTTTCCGTCTCCGGCTTAGCAA 1027
Qy   301 CCCCACATTTGTTGCTTTAGAGTCATTAATCCAACGAATCCCATTAGCGTTGATGCACTAA 360
Db   1026 CAAATGCAATTTACCCTCAATATA------AAGCCCAGCCAATGCAGCGAACTGTACT 976
Qy   361 GCCCATGTTTGTTACCTTTTGAATCATTAAATATGTCCTTGCTACTATCAATATCATCAATATA 420
Db   975 CCCCACATTTTGTTGCTTTAGAGTCATTAATCCAACGAATCCCATTAGCGTTGATGCACTAA 916
Qy   421 TTCACAGCGATGCGGTAGCCCTTTAGAGTCATTAAGGGTATCAAGCATGGCATCCCATGGG 480
Db   915 TTGAAAACGATGACTAACCCTTTGAAATGACGAAGTGCGGGTACAAATTGAATCTAAATT 856
Qy   481 AAGATTGGCAGCTGTGCCTAGGCAAGTGCAGATAAGGCATTAGTAGGTTATGCTTGCC 540
Db   855 AATACCTATAGCTTGTGCCAATGCTGCGCCAAAATGTTCATATAATTATATGGCGACC 796
Qy   541 TTTGATTTTTAATTGATGGATTGGTAAAAAGTTTGTTTTTTTTGATGATAATAATGCATACC 600
Db   795 AACCAAATGTAGCTTCTTCACAAGGTAAAATCACTTCATCTTTACCATTAAATATATGCTT 736
Qy   601 ATCACCATCAAAATAAATAAAAAAATCTGC-CGTTTGGTGTGGCCGTAAGTAAGCTATTGTTG 659
Db   735 GCCATTTCAGTTTTAGCCAATAATCCGGCAACACCTACTTTAGCATTATGATAAGTGTTT 676
Qy   660 TGATAATGGTGTGATTGGGGGGCACACCTTGTCAGTGTCAAGCATTGCTTGGCTAAAT 719
Db   675 CGCTTGATTTTCGTTTTCCCCAAAAGTCAGCC------------- 644
Qy   720 TGGCATCGTCTTGGTAGATGATGGGGTTTGGCAATTTCAAAAATACCCAATTTTTGTG 779
Db   643 ---TATCTTCATTGTTCAACACCTACTTTAGCATTATGATAAATGCGTAATTTGCTT 587
Qy   780 CCAAATAATCCACCATGCCGATCAGTCGCCGATCAAGATGATCAGCGGAGAGATTTAAAAATTG 839
Db   586 GGCCGATAATCTTCTAAATCATATAGCGATCCAATAATTTTCAGTCAGTGTCAAGACAG 527
Qy   840 TTGCCCCTTGAGCTCGCCAAATTTGAGATATCTCAAGCTGAAACAACTTGACAGCTCAAGCA 899
Db   526 TCGCTGCTGCCCGCAGCTTTTAAGCTATAAGGTTGTCTCAAGCTGAAAAAGCTCTAGTA 467
Qy   900 CCACCAAAATCCAATATTTTTGGATCGTTAATAATTCAAGTGGGTGCCAAAAATTACCGC 959
Db   466 CATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAATATTTTTCCGC 410
Qy   960 CGACACCAACAAATCATGCCTGCCATGTTTTCCCATCTCGCCTACTAAGTCGTAACAGTGC 1019
Db   409 CCATACCAACTTTCACACCAGCAGCTTTCGCCATTTCATTAAAACTAAAGTAGTTACGGTAC 350
Qy   1020 TTTTCGGCATTTGAGCCTGTGAGTGGCGGATGATTGGTGTGTTTTGTTCGCTTGAGTTAGATCTC 1079
Db   349 TTTTTACCATTTTGAACCTGTAATCCCACAATTGCCGTCGCGCGGCCA------ 298
Qy   1080 GTGCTTTGAGTGTATCAACAATAGAACAATTGAACATACTCGCACAGGAATGCCCTGCGGGCCA- 1139
Db   297 ------GAATAATTCAATATCGCGATTACTTCCACTCCCGCTTTAA 257
Qy   1140 TGGGCGCGGACGAATACTTCGGGGTTTGAATTTCTGTGTTTTTACCGCAAGCCCTGGGCTCCCAA 1199
Db   256 GTGCGGGTTTGAATTTCACTCGTGTTATTTTACGCAGCCCTGGGCTAATGCCGATTGAATCTGATCGG 197
Qy   1200 CTGACTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAGCCCTGTGGGCCAGTT 1259
Db   196 TTTCAAGTAACCATTCCTGATTTAAACTACACCAGGGGATTATTTGAGGNAGTT 137
Qy   1260 TATCTGCCAAGGTGATCGGATTGGGGATATTACCATCGTCTAACTTTGTGACCGTGATTGA 1319
Db   136 TATCAATACCAGTAGGGATTTTTTCGAGTATCGAAATCACACGGAATATTAGCCTGTTGGGATA 77
Qy   1320 CCAAAAAATTCACAGCTGACGCAGTCCGAGAGCTGCCAAACCGACGGATCCAGCCGATGATTAGTAATATATTTTTT 1377
Db   76 AGAGAATAATCCACACAAGAAAGACCTGTTTTGCAAGCCGATGATTAGTAATATATTTTTT 19

```
Db   76 AGAGATAATCCACACAAGAAGACCTGTTTTGCCAAGCCCGATGATAGTAATATATTTTT   19
```

```
Qy 1086 TTGATTAATTAATTCAGCTAAATTCTGAAAAATCAGCCCCTTTTCCGTCTCCGCCTAGCAA 1027
Db  301 CAATGCAATTACCCTCAATATA--------AAGCCCAGCCAATGCAGCCAACTGTACT  360
Qy 1026 CAAATGCAATTACCCTCAATATA--------AAGCCCAGCCAATGCAGCCAACTGTACT  976
Db  361 GCCGATGTTTGTACCTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA  420
Qy  975 CCCCACATTTGTTGTCTTTAGAGTCATTAATCAACGAAATGCCATTAGTCTTGACTAA  916
Db  421 AAGATTGGCAGCTGTGCCTAAGGCAAGTCGCAGATAAGGCATTGAGTTATGTTATGCC  480
Qy  915 TTGAAAACGATGATCTAACCCTTTGAAAATGACGAAGGTGCGGGTACCGAAATTGAAATT  856
Db  481 AAGATTGGCAGCTGTGCCTAAGGCAAGTCGCAGATAAGGCATTGAGTTATGTTATGCC  540
Qy  855 AATACCTATAGCTTGTGCCAAATGCTGTTGCTGCCAAAATCTTCATAATAATTATGCGACC  796
Db  541 TTTGATTTTTAATTGATTGGATTGGTAAAAGTTTGTTTTTGATGATATAATGCCATACC  600
Qy  795 AACCAAATGTAGCTTCTTCACACGCTAAATCACTTTACCATTAAAATATTGCTT  736
Db  601 ATCACCATCATCAAATAAATAAAAATCTGC-CGTTTGGTGTGGCTAAGTAAGCTATTTGTTG  659
Qy  735 GCCATTTTCAGTTTTTTAGCCAATAATCCGGACTATATTTCCGCAAAAGAAACGGTATGTTTT  676
Db  660 TGATAAATGGCTGTGATTTGGGGCAACGCTTGTCAGTGTGGTGTCAACCATTGCTTGGCTAAAT  719
Qy  660 TGATAAATGGCTGTGATTTGGGGCAACGCTTGTCAGTGTGGTGTCAACCATTGCTTGGCTAAAT  719
Db  675 CGCTTGATTTTGTTTTTCCCAAAAGTCAGCC------------  644
Qy  675 CGCTTGATTTTGTTTTTCCCAAAAGTCAGCC------------  644
Db  720 TGGCATCGTGTTCTTGGCTAGATGATGGGCGGTTTGGCCAATTTTTCAAAAATACCGCAATTTTGTG  779
Qy  720 TGGCATCGTGTTCTTGGCTAGATGATGGGCGGTTTGGCCAATTTTTCAAAAATACCGCAATTTTGTG  779
Db  643 ---TATCTTCATTGTTCACACACTACTTTAGCGATTATAAATGCGTAATTTTGCTT  587
Qy  643 ---TATCTTCATTGTTCACACACTACTTTAGCGATTATAAATGCGTAATTTTGCTT  587
Db  780 CCAAATAATCCACCATGCCATCGTCGTCAAGATGATCAGGCGGAGAGATTAAAATTG  839
Qy  780 CCAAATAATCCACCATGCCATCGTCGTCAAGATGATCAGGCGGAGAGATTAAAATTG  839
Db  586 GGCGATAATCTTCTPAAATCCATATAGCGATCCCATATGATCTTCAGTCACGTTCAAGACAG  527
Qy  586 GGCGATAATCTTCTPAAATCCATATAGCGATCCCATATGATCTTCAGTCACGTTCAAGACAG  527
Db  840 TTGCCCCTTGAGCTGCCAAATTTGAGATATGCCAAGCTCAAGCTCAAGCA  899
Qy  840 TTGCCCCTTGAGCTGCCAAATTTGAGATATGCCAAGCTCAAGCTCAAGCA  899
Db  526 TCGCTGCCGCGGAGCTTTTAAGCTATAAGTTGTCTCAAGCTGAAAAGCTCTAGTA  467
Qy  526 TCGCTGCCGCGGAGCTTTTAAGCTATAAGTTGTCTCAAGCTGAAAAGCTCTAGTA  467
Db  900 CCACCAAAATCCATATATTTGGATCGTTAATAATTCAAGTGCGAGGCGGCTGCCAATAATTACCGC  959
Qy  900 CCACCAAAATCCATATATTTGGATCGTTAATAATTCAAGTGCGAGGCGGCTGCCAATAATTACCGC  959
Db  466 CATAAAGTTCACAAACTTCATTC---AACAATGACAAAGCGGGAATCCCAATATTCCGC  410
Qy  466 CATAAAGTTCACAAACTTCATTC---AACAATGACAAAGCGGGAATCCCAATATTCCGC  410
Db  960 CGACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGC 1019
Qy  960 CGACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGC 1019
Db  409 CCATCAACTTTCACACCAGCAGCTTTCGCCCATTTCATAAACTAAAGTAGTTACGGGTAC  350
Qy  409 CCATCAACTTTCACACCAGCAGCTTTCGCCCATTTCATAAACTAAAGTAGTTACGGGTAC  350
Db 1020 TTTTGGCATTTGAGCCTGTGATGGGGGATGAAATGGTGTTGTGTTGCTTGAGTTAGATCTC 1079
Qy 1020 TTTTGGCATTTGAGCCTGTGATGGGGGATGAAATGGTGTTGTGTTGCTTGAGTTAGATCTC 1079
Db  349 TTTTTACCATTTGAACCTGTAATCCCCACAATTGGCTTTGTCGCTGGCGGCA  298
Qy  349 TTTTTACCATTTGAACCTGTAATCCCCACAATTGGCTTTGTCGCTGGCGGCA  298
Db 1080 GTGCTTTGAGTGTATCAATGAACATCGGTGTGATCACAGGAATGCCTTGAGCTT 1139
Qy 1080 GTGCTTTGAGTGTATCAATGAACATCGGTGTGATCACAGGAATGCCTTGAGCTT 1139
Db  297 ------------GAATAATTCAATATATCCCGATTACTTCCACTCCCGCTTTAA  257
Qy  297 ------------GAATAATTCAATATATCCCGATTACTTCCACTCCCGCTTTAA  257
Db 1140 TGGCGGCGGACAATACTCGGCCGTTCTTGGGTTAATGCGGGGGCTGAGTGATGATCTCGG 1199
Qy 1140 TGGCGGCGGACAATACTCGGCCGTTCTTGGGTTAATGCGGGGGCTGAGTGATGATCTCGG 1199
Db  256 GTGCGGTTTGAATTTCTGGTGTTTTTACCGGCTAATAACAATCATATCGC  197
Qy  256 GTGCGGTTTGAATTTCTGGTGTTTTTACCGGCTAATAACAATCATATCGC  197
Db 1200 CTGAGCTTAGAGTTCACTGTGCCAAGCTGCCAAAATGACAACGACGTGTGGGCAGTT 1259
Qy 1200 CTGAGCTTAGAGTTCACTGTGCCAAGCTGCCAAAATGACAACGACGTGTGGGCAGTT 1259
Db  196 TTTCAAGTAACCATTCCTGATTTAAAACTACCAGTTAAGAGGATATTTGAGGAAGTT  137
Qy  196 TTTCAAGTAACCATTCCTGATTTAAAACTACCAGTTAAGAGGATATTTGAGGAAGTT  137
Db 1260 TATCTGCCAAGGTAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGATTGA 1319
Qy 1260 TATCTGCCAAGGTAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGATTGA 1319
Db  136 TATCAAATACCAGTAGGATTTTTTCGAGTATCAAATCACACGAATATTAGCCTGTCGGGATA   77
Qy  136 TATCAAATACCAGTAGGATTTTTTCGAGTATCAAATCACACGAATATTAGCCTGTCGGGATA   77
Db 1320 CCAAAAAATTCACAGCTGACAGTCGCCCAAACGACGACGGGCATATTTTT 1377
Qy 1320 CCAAAAAATTCACAGCTGACAGTCGCCCAAACGACGACGGGCATATTTTT 1377
```

Query Match 11.6%; Score 160.4; DB 10; Length 1314;
Best Local Similarity 49.8%; Pred. No. 2.1e-33;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```
Qy    1 TAAAGGGTGAACCATGTTAACAAAACTTATCACCTGGCTCATTATATATTTTAAATTGATC   60
Db 1308 TAAACATTGAGCTAAATGCGTAAATTCTTCGCGCGGCTTTTCAAAAGAGCAAAACTGAATC 1249
```

```
Qy          61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 120
                |  || ||||||||||| ||||| |  ||              | |  | |  |  |||
Db        1248 GAGACTTGCACAAGCAGGCGACAATAATACCA------------TATCTCCGCTTTGCAA 1201

Qy         121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 180
               |  ||  ||   || |       | || |  ||| ||     |||       || || ||
Db        1200 TGTTGGGCGTAAAAATTCTATCGCTTGTTCCATTGTATCGAACAA------ATAACTTTG 1147

Qy         181 GAGTAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACTATCCACCCA 240
                  | ||  ||   |||  | ||   ||||| ||||| |  ||||   | ||
Db        1146 CGATGAAAATTTTGCAAGCAGCGCACCATCTCGACCAAAACAATAACAAATAATGTGTGG 1087

Qy         241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
                ||  |||   || | |  |  ||   ||||||    ||||| ||    || || |  |
Db        1086 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCCGTCTCCGCCTAGCAA 1027

Qy         301 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
               |||   || |           |||         ||||||| | | ||| ||||  ||
Db        1026 CAAATGCAATTTACCCTCAATATA---------AAGCCCAGCCAATGCAGCAACTGTACT 976

Qy         361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
                || |  |||||  |||| || |||||||   |     ||| || |   | ||  |   |
Db         975 CCCCACATTTGTTGCTTTAGAGTCATTAATCCAACGAATGCCATTAGCTTGATGCACTAA 916

Qy         421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCATGGCATCCATGGG 480
               || | | |||||   || ||||||||||     ||| |    | | || | ||| |
Db         915 TTGAAAACGATGATCTAACCCTTTGAAATGACGAAGTGCGGTACGAATTGAATCTAAATT 856

Qy         481 AAGATTGGCAGCTGTGCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
               || |     ||||    | || ||   ||| |  ||    || |     |||||    ||
Db         855 AATACCTATAGCTTGTGCCAATGCTGTTGCTGCCAAAATGTTCATATAATTATGGCGACC 796

Qy         541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTTGTTTTTTTGATGATATAATGCCATACC 600
                   | | |   ||  |     |||||||     |  | ||       |||
Db         795 AACCAATGTAGCTTCTTCACAAGGTAAAATCACTTCATCTTTTACCATTAAATATTGCTT 736

Qy         601 ATCACCATCATCAAATAAATAAAAATCTGC-CGTTTGGTGGCTAAGTAAGCTATTTGTTG 659
                 ||   |||     ||   || |||| || |  ||    || ||  || |  | ||||
Db         735 GCCATTTTCAGTTTTTAGCCAATAATCCGCACTATTTTCCGCAAAAGAAACGGTATGTTT 676

Qy         660 TGATAATGGTGTGATTTGGGGCAACGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAAT 719
                | |     |  | |||     || |   | |
Db         675 CGCTTGATTTTCGTTTTCCCCAAAAGTCAGCC---------------------------- 644

Qy         720 TGGCATCGTCTTGGTAGATGATGGCGGTTTGGCAATTTTCAAAAATACGCAATTTTTGTG 779
                   ||| || | ||  |  |   |   ||    ||| | | |||| || ||||||  |
Db         643 ---TATCTTCATTGTTCAACACACCTACTTTAGCATTATGATAAATGCGTAATTTTGCTT 587

Qy         780 CCAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTG 839
                   ||||||  | |   |    |  ||||| | ||||||  | |  |  || || |  |
Db         586 GGCGATAATCTTCTAAATCCATATAGCGATCCATATGATCTTCAGTCACGTTCAAGACAG 527

Qy         840 TTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCA 899
               | ||  |   ||||   ||  |  |  |   ||||||||||||||| || ||||| || |
Db         526 TCGCTGCCGCAGCTTTTAAGCTATAAGTTGTCTCAAGCTGAAAACTAGAAAGCTCTAGTA 467

Qy         900 CCACCAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTACCGC 959
               |    |  ||   || ||   ||   || |||   |  || ||  | |||||||| ||||
Db         466 CATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAATATTTCCGC 410

Qy         960 CGACACCAACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTAATGTCGTAACAGTGC 1019
               | | ||||||  |||  || |||  ||| |||| ||    ||||| || || || || |
Db         409 CCATACCAACTTTCACACCAGCAGCTTTCGCCATTTCATAAACTAAAGTAGTTACGGTAC 350

Qy        1020 TTTTGGCATTTGAGCCTGTGATGGCGATGATTGGTGTGTTTGTTGCTTGAGTTAGATCTC 1079
               ||||  |||||||| ||||| ||  | |  |||||  |  | | |||  |
Db         349 TTTTACCATTTGAACCTGTAATCCCCACAATTGGCTTTGTCGCTGCGCGGCA-------- 298

Qy        1080 GTGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAATGCCTTGAGCTT 1139
                                  ||| |||| || ||||| ||| ||    |  ||
Db         297 -------------------GAATAATTCAATATCGCCGATTACTTCCACTCCCGCTTTAA 257

Qy        1140 TGGCGGCGACAATACTCGGCGTTCTTGGGTTAATGCCGGGGCTGATGATGATCTGATCGG 1199
                 ||||    |||    || ||| ||     ||  || ||||| || |  |||  ||||
Db         256 GTGCGGTTTGAATTTCTGGTGTTTTTACCGCAAGCCCTGGGCTAATAACAATCATATCGC 197

Qy        1200 CTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGCAGTT 1259
                |     |||   |||    |  || || |||  ||||  | | |    |   || ||||
Db         196 TTTCAAGTAACCATTCCTGATTTAAACTACCAGTATGAAGAGGGATATTTTGAGGAAGTT 137

Qy        1260 TATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACCGTGATTGA 1319
               ||||        ||||||||      |  ||| || |  | ||  ||       ||    |
Db         136 TATCAATACCAGTAGGATTTTTTCGAGTATCAATCACACGAATATTAGCCTGTTGGGATA 77

Qy        1320 CCAAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGACGGCATATTTTTT 1377
                 | | ||| ||||   || || || |   |||| |  ||||| ||  | |   |||||
Db          76 AGAGATAATCCACACAAGAAAGACCTGTTTTGCCAAGCCCGATGATAGTAATATTTTT 19

RESULT 7
US-11-194-246-220/c
; Sequence 220, Application US/11194246
; Publication No. US20050272089A1
; GENERAL INFORMATION:
;  APPLICANT: Mott, John
;  APPLICANT:  Trepod, Catherine
;  APPLICANT:  Arvidson, Staffan
;  TITLE OF INVENTION: CRITICAL GENES AND POLYPEPTIDES OF HAEMOPHILUS INFLUENZAE AND METH
;  TITLE OF INVENTION:  USE
;  FILE REFERENCE: 00592.US1 (M&R 268.05920101)
;  CURRENT APPLICATION NUMBER: US/11/194,246
;  CURRENT FILING DATE:  2005-08-01
;  PRIOR APPLICATION NUMBER: US/10/274,586
;  PRIOR FILING DATE: 2002-10-21
;  PRIOR APPLICATION NUMBER: US 60/345,438
;  PRIOR FILING DATE: 2001-10-19
;  NUMBER OF SEQ ID NOS: 621
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 220
;   LENGTH: 2925
;   TYPE: DNA
;   ORGANISM: ARTIFICIAL
;   FEATURE:
;   OTHER INFORMATION: Sequence of the murD coding sequence and flanking regions.
US-11-194-246-220

  Query Match             11.6%;  Score 160.4;  DB 15;  Length 2925;
  Best Local Similarity   49.6%;  Pred. No. 3.4e-33;
  Matches  683;  Conservative    0;  Mismatches  606;  Indels   89;  Gaps    7;

Qy           1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTTAAATTGATC 60
               ||||   ||| | |  |    ||| |  || || ||||  | | ||       ||| |||||
Db        2240 TAAACATTGAGCTAAATGCGTAAATTCTTCGCCGCGCTTTTCAAAAGAAGCAAACTGATC 2181

Qy          61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGCTGCCAA 120
                |  || ||||||||||| ||||| |  ||              | |  | |  |  |||
Db        2180 GAGACTTGCACAAGCAGGCGACAATAATACCA------------TATCTCCGCTTTGCAA 2133

Qy         121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 180
               |  ||  ||   || |       | || |  ||| ||     |||       || || ||
Db        2132 TGTTGGGCGTAAAAATTCTATCGCTTGTTCCATTGTATCGAACAA------ATAACTTTG 2079

Qy         181 GAGTAAATCTTGCTCAATTATCGGGGCATCTTGACCGATTAAATACACACTATCCACCCA 240
                  | ||  ||   |||  | ||   ||||| ||||| |  ||||   | ||
Db        2078 CGATGAAAATTTTGCAAGCAGCGCACCATCTCGACCAAAACAATAACAAATAATGTGTGG 2019

Qy         241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
                ||  |||   || | |  |  ||   ||||||    ||||| ||    || || |  |
Db        2018 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCCGTCTCCGCCTAGCAA 1959

Qy         301 CAATGCCAGTGATGATATGCCATAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
               |||   || |           |||         ||||||| | | ||| ||||  ||
Db        1958 CAAATGCAATTTACCCTCAATATA---------AAGCCCAGCCAATGCAGCAACTGTACT 1908

Qy         361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCAATATCATCAATATA 420
                || |  |||||  |||| || |||||||   |     ||| || |   | ||  |   |
```

RESULT 8

; TITLE OF INVENTION: Antisense Antibacterial Cell Division
; TITLE OF INVENTION: Composition and Method
; FILE REFERENCE: 0450-0033.30
; CURRENT APPLICATION NUMBER: US/09/754,468
; CURRENT FILING DATE:
; PRIOR APPLICATION NUMBER: 2001-01-04
; PRIOR FILING DATE: 2000-01-04
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 44
; LENGTH: 11498
; TYPE: DNA
; ORGANISM: Hemophilis influenza
US-09-754-468-44

Query Match   11.6%;   Score 160.4;   DB 4;   Length 11498;
Best Local Similarity   49.6%;   Pred. No. 7.5e-33;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

[DNA sequence alignment data in two columns — Query (Qy) and Database (Db) sequences with alignment match bars]

```
Db  8400  GGCGATAATCTTCTAAATCCATATAGGCGATCCATAATGATCTTCAGTCAGTTCAAGACAG  8341

Qy   840  TTGCCCTTGAGCTCCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA   899
Db  8340  TCGCTGCCCGCAGCTTTAAGCTATAAGTTGTGTCTCAAGCTGAAAAACTTAGAAAAGCTCTAGTA  8281

Qy   900  CCACCAAAATCCATATTTTGGATCGTTAATAATTCAAGTGCAGGCGTGCCAATATTATACCGC   959
Db  8280  CATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAATATTTTCCGC  8224

Qy   960  CGACACCAAACAAATCATGCCTTGCCATGTTTTTGCCATCTCGCCTACTTAATGTCGTTAACAGTGC  1019
Db  8223  CCAATACCAACTTTCACACCACGAGCTTTCCCCATTTCATAAACTAAAGTAGTTACGGTAC  8164

Qy  1020  TTTTGGCATTTGAGCCTGTGATGAGGCCGATGAGTTGGGTCGTGTTTGTTTGCTTGAGTTAGATCTC  1079
Db  8163  TTTTACCATTTGAACCTTGTTAATCCCCACAATTGGCTTTGTCGCGTCGGCGGCA------  8112

Qy  1080  GTGCTTTGAGTGTATCAAATGAACAATTGAACATCGCTGATCACAGGAATCCCTTGAGCTT  1139
Db  8111  ----------------GAATAATTCAATATATCGGCGATTACTTCCACTCCCGCTTTAA  8071

Qy  1140  TGGGCGGCGGACAATACTCGGCGGTTCTTGGGTTAATCGCGGGGCTGATGATGATCTGATCGG  1199
Db  8070  GTGCGGTTTGAATTTCTCGTGTTTTACCGCAGGCCCTGAATAACAATCATATCGC  8011

Qy  1200  CTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGACAGCCTGTGGGCAGTT  1259
Db  8010  TTTCAAGTAACCATTCCTGATTTAAACTACCAGTATGATGAAGGGGATATTTTGAGGAAGTT  7951

Qy  1260  TATCTGCCAAGGTAGGGATTGGGATTTACCATCGATGATGATGTGACCGTGATTGA  1319
Db  7950  TATCAATACCAGTAGGATTTTTTTCGAGTATCAAATCACACGAATAATTAGCCTGTTCGGGATA  7891

Qy  1320  CCAAAAAAATTCACAGCTGACAGTCCCAAAACCGACGACGGGCCCATATTTTTT  1377
Db  7890  AGAGATAACCCACACAAGAAAAGACCTGTTTTGCCAAGCCCGATGAATAGTAATATATTTTT  7833
```

SEQ ID NO 1
  LENGTH: 1830121
  TYPE: DNA
  ORGANISM: Haemophilus influenzae
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (4747)..(4747)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (9921)..(9921)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (10150)..(10150)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (29298)..(29298)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (36543)..(36543)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (36551)..(36551)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (36636)..(36636)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (40808)..(40810)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (44416)..(44416)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (44905)..(44905)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (44975)..(44975)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (45593)..(45593)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (45732)..(45732)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (47036)..(47036)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (51334)..(51334)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (51602)..(51602)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (51786)..(51786)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (51805)..(51805)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (55369)..(55369)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (65309)..(65309)
    OTHER INFORMATION: n equals a, t, g or c
  FEATURE:
    NAME/KEY: misc_feature
    LOCATION: (65313)..(65313)
    OTHER INFORMATION: n equals a, t, g or c

FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:

NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)

Query Match 11.6%; Score 160.4; DB 8; Length 1830121;
Best Local Similarity   49.6%; Pred. No. 1.4e-31;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```
Qy         1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTAAATTGATC 60
Db   1204439 TAAACATTGAGCTAAATGCGTAAATTCTTCGCGCGGCTTTTCAAAAGAGGAAACTGATC 1204380

Qy        61 CATGTCTAGCACAAAGCAGGTGACAAAAGCACAGCCCTTAGCGCGTGCTTTGACTGGCTGCCAA 120
Db   1204379 GAGACTTGCACAAGCAGGCCGACAAATAATACCA------TATCTCCGCTTTGCAA 1204332

Qy       121 TACTGCCCGATCAAGTGTACCAAACCTGATATATTTTAGATGCCAAATGTGCATCACCTTT 180
Db   1204331 TGTTCGGGCGTAAAAAATTCTATCGCTTGTTCCATTGTATCCGAACAA------ATAACTTTG 1204278

Qy       181 GAGTAAAATCTTGCTCAATTATCGGGGCCATCTTGACCGGATTAAAATACACTATCCACCA 240
Db   1204277 CGATGAAAATTTTGCAAGCAGCGCGCACCCATCTCGACCAAAACAAATAACAAATAATGTGTGG 1204218

Qy       241 CTGCATAACAACATATGCTGACAGCCTCACTAAAAATCTTGACCTTTACCAAGCCCACCAAAAT 300
Db   1204217 TTGATTAATTAATTCAGCTAAATTCTGAAAAAATCAGCCCCTTTTCCGTCTCCGCCTAGCAA 1204158

Qy       301 CAATGCCCAGTGATGATGATGATATGCCATAAAACGTGCCAAGCCCATCAATTGCTCGCAATGGTTTGA 360
Db   1204157 CAAATGCAATTTACCCTCAAAATA-------AGCCCAGCCAATGCAGCAAACTGTACT 1204107

Qy       361 GCCGATGTTTGTTGTACCTTTTGAATCATTAAAATATGCCTTGGCTTACTCAAATATCATCAAATATA 420
Db   1204106 CCCCACATTTGTTGCTTTAGAGTCATTAATCCAACGAATGCCATTAGCTTGATGCGACTAA 1204047

Qy       421 TTCACAGCGATGGCGGTAGCCCTTTTGAAAGTCTTAAGGGTATCAAGCATGCCATCCATGGG 480
Db   1204046 TTGAAAACGATGATCTAACCCTTTGAAAATGACGAAGTGCCGGTACGAAATTGAATCTAAATT 1203987

Qy       481 AAGAATTGGGCAGCTGTGTCCTAAGGCAAGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCC 540
Db   1203986 AATACCTTATACGTTGTGCCAATGCTGTTGCCTGCAAAAATGTTCATTATAATTATGGGGACC 1203927

Qy       541 TTTGATTTTTTAATTGATGGGATTGGTAAAGTGTTTGTTTTTTTTTTTGTGATGATTATTATGCCATACC 600
Db   1203926 AACCAATGTAGCTTCTTCACAAGGTAAAAAATCTGC-CGTTTGGTTGGCTAAGTAAGCTATTTGTTG 1203867

Qy       601 ATCACCCATCATCAAATAAAATAAAAAATCTGC-CGTTTGGTTGGCTAAGTAAGCTATTTGTTG 659
Db   1203866 GCCCATTTTCAGTTTTTTAGCCAAATAATCCGCACTATTTCCGCAAAAAGAAAACGGGTATGTTT 1203807

Qy       660 TGATAATGGTGTGGATTGGGGGCAACGCCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAAT 719
```

```
Qy    720  TGGCATCGTCTTGGTGATGATGGCGGTTTGGCAATTTCAATAAATACGCGAATTTTTGTG  779
                |||                                                          |||
Db  1203774  ---TATCTTCATTGTTCAACACACTTACTTTAGCATTATGATAAATGCCGTAATTTGCTTT  1203718

Qy    780  CCAAATAATCCACCATGCCATCGTGCCCGATCAAGATGATCAGCGGGAGAGAGATTTAAAATTG  839
           |||||
Db  1203717  GGGCGATAATCTTCTAAATCCATATAGCGATCCATATGAICTTCAGTCACGTTCAAGACAG  1203658

Qy    840  TTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGACAGCTCAAGCA  899
                                         |||
Db  1203657  TCGGCTGCCGCAGCTTTTAAGCTATAAGTGTGTCTCAAGCTGAAAAACTTCAAGACAGTA  1203598

Qy    900  CCACCAAATCCAATATTTTGGATCGTTAATTAATTCAAGTGCAGCCGTGCCAATATTACCGC  959
           |||||||                                                     |||
Db  1203597  CATAAAGTTCACAATCTTCATTC---AACAATGACAAAGCGGGAATCCCAATATTTCCGC  1203541

Qy    960  CGACACCAACAATCATGCCTGCGCATGTTTTGGCCATCTCGCCTACTTAATGTCGTAACAGTGC  1019
                                    |||
Db  1203540  CCATACCAACTTTCACACCAGCAGCTTTCGCCCATTCATTAAAACTAAAGTAGTTACGGTAC  1203481

Qy   1020  TTTTTGGCATTTGAGCCTGTGATGGCGGATGGTGGTTGTGGTTTGTTGCTTGAGTTAGATCTC  1079
                                         |||
Db  1203480  TTTTACCATTTGAACCTGTAATCCCCACAATTGCTTGTTGTCGCTGCGCGGCA-------  1203429

Qy   1080  GTGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGCTGATCACAGGAATGCCTTGAGCGTT  1139
                                         |||
Db  1203428  -------------GAATRATTCAATATGCGCCGATTACTTCCACTCCCGCTTTAA  1203388

Qy   1140  TGGCGGGCGACAAATACTCCGCGTTCTTGGGTTAATGCCCGGGGGCTCGATGATGGATCGGATCGG  1199
           |||                                                          |||
Db  1203387  GTGCGGGTTTGAATTTCTGGTGTTTTTACCCGCAAGCCCTGGGCTAATAACAATCATATCGC  1203328

Qy   1200  CTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAATGAACGCCTGTGGGGCAGTT  1259
                                         |||
Db  1203327  TTTCAAGTAACCATTCCTGATTTAAACTACCAGTATGTGAAAGAGGGGATATTTTGAGGAAGTT  1203268

Qy   1260  TATCTGCCAAGGTAGGATTGGGATTACCATCGATCATGATGCTAACTTTGTGACCGTGATTGA  1319
           |||                                                          |||
Db  1203267  TATCAAATACCAGTAGGATTTTTCGAGTATCAAATCACCGAAATATTAGCCGTTGGGGATA  1203208

Qy   1320  CCAAAAATTCACAGCTGACAGTCCAGAGCTGCCCAAACCGACGGCGCCATATATTTTTT  1377
                                         |||
Db  1203207  AGAGATAATCCACACAAGAAAGACCTGTTTTGCCAAGCCCGATCGATAGTAATATTTTT  1203150
```

FEATURE:
NAME/KEY: misc_feature
LOCATION: (9921)..(9921)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (10150)..(10150)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (29298)..(29298)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36543)..(36543)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36551)..(36551)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (36636)..(36636)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (40808)..(40810)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44416)..(44416)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44975)..(44975)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (44905)..(44905)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45593)..(45593)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (45732)..(45732)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (47036)..(47036)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51334)..(51334)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51602)..(51602)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g

FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122167)..(122167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152530)..(152530)

Query Match   11.6%; Score 160.4; DB 9; Length 1830121;
Best Local Similarity   49.6%; Pred. No. 1.4e-31;
Matches 683; Conservative   0; Mismatches 606; Indels   89; Gaps   7;

```
Qy         1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATAATTTTAAATTGATC 60
             ||| |  ||| | |   |  || |  ||| | | ||||| |||| |||  ||||||||
Db   1204439 TAAACAATTGAGCTAAATGCGTAAATTCTTCGCCCGGCTTTTCAAAAGAGCAAAACTGATC 1204380

Qy        61 CATGCTTAGCACAAGCAGGGTGACAAAAGCACAGCCCTAGCGGGTGCTTGACTGGGCTGCCAA 120
             |||   ||||| | |  |  |   | ||||  ||| |||||  ||||||||  |  |  |||||
Db   1204379 GAGACTTGCACAAGCAGGGGGACAATAATACCA------TATCTCCGCTTTGCAA 1204332

Qy       121 TACTCCCGATCAAGTGTACCAACCTGATATATTTTAGATGATATGTTGCATCACCTTT 180
             |  | ||| | |   |||| | |||| | | ||  | | |  |    |  ||||||||
Db   1204331 TGTTGGGCGTAAAAATTCTATCGCTTGTTCCATTGTATCGAACAA------ATAACTTTG 1204278

Qy       181 GAGTAAATCTTGCTCAAATTATCGGGGGCCATCTTGACCGGATTAAAATACACACTATCCACCCA 240
             | || || |||||  |  | | || | ||||||| |||| | ||||||  | || |   || |||
Db   1204277 CGATGAAAAATTTTGCAAGCAGCGGCACCATCTCGACCAAAACAAATAAACAAATAAGTGTGG 1204218

Qy       241 CTGCATAACATATGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAT 300
             ||| || |   || ||  |  || | ||||||||  ||||  |||||| ||| |||| |
Db   1204217 TTGATTAATTAATTCAGCTAATTCTGAAAAAATCAGCCCCCTTTTCCGTCTCCGGCCTAGCAA 1204158

Qy       301 CAATCGCCAGTGATGATAATGCCATAAAACCGTGCCAAGCCCATCAATTGCTCCAATGGTTGA 360
             |||| || | | |  |||| || | ||||||| || ||||| | ||| | | | ||||  ||
Db   1204157 CAAATGCAAATTACCCTCAATATA---------AAGCCCAGCCAAATGCAGCAACTGTACT 1204107

Qy       361 GCCGATGTTTGTACCTTTTGAATCATTAAAATATGCCTTGCTATCATCAATATCACAATATA 420
             |||||||| ||| | |||| | |||| | || || | | | |||||||||| ||| |||||
Db   1204106 CCCCACATTTGTTCGCTTTAGAGTCATCATTAATCCAACGAATCGCCATTAGCTTGATGCGACTAA 1204047

Qy       421 TTCACAGCGATGCGCGGGTAGCCCCTTGAAAGTCTTAAGGGGTATCAAGCACATGGCCATCCATGGG 480
             || || |||||| | || ||||| ||||||| | ||| ||| | | |||  |  ||| | |||||
Db   1204046 TTGAAAACGATGATCTAACCCTTTGAAAATGACGAAGTGCGGGTACGAAATTGAAATCTAAAATT 1203987

Qy       481 AAGATTGGCAGCTGTGCCTAAGGCAATGTGCAGATAAAGGACATTGAGTAGGTTATGCTTGCC 540
             || | ||||| ||||||||||  ||| || | |||||  | |||| | |||| ||| ||| ||
Db   1203986 AATACCTATAGCTTCTGTGCCAAATGTTCATATATAATTATGCGACC 1203927

Qy       541 TTTGATTTTAAATTGATGGATTGGTAAAAGTTTGTTTTTTGATGATATAAATGCCATACC 600
             ||||| || || ||||||||| ||| |||| ||||||||||||||||||||||||||| ||
```

```
Qy    601  ATCACCATCATCAAAATAAAATAAAAATCTGC--CGTTTGGTGGCTAAGTAAGCTATTTGTTG  659
           |||||||||||||||||||||||||||||||||  |||||||||||||||||||||||||||||
Db 1203926  AACAAATGTAGCTTCTTCACAAGGTAAAATCACTTCATCTTTTACCATTAAAATATTGCTT  1203867

Qy    660  TGATAATGGGTGTGATTTGGGGGCAAGCCTTGTTCAGTGTGGTCAAGCATTGCTTGGCTAAAT  719
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db 1203866  GCCATTTTCAGTTTTTTAGCCAATAATCCGACTATTTTCCCGCAAAAAGAAACGGTATGTTT  1203807

Qy    720  TGGCATGCGTCTTGGTAGATGATGGGCCGTTTGGGCAATTTCAAAAAATACGCAATTTTTGTG  779
                 |||||||||||||||||||||||
Db 1203806  CGCTTCGATTTTCGTTTTTCCCAAAAGTCAGCC------------------------  1203775

Qy    780  CCAAATAATCCACCATGCCATCGTGCCGATCAAGATGATCAGCGGAGAGATTTAAAATTG  839
           |||||||||||||||||||||||
Db 1203774  ---TATCTTCATTGTTCAACACACTTTAGCATTATATGATAAATCGGTAAATTCTTT  1203718

Qy    840  TTGCCCCTTGAGCTCGCCAAAATTTGAGATATGCTCAAAGCTCAAAACTTGACAGCTCAAGCA  899
           ||||||||||||||||||||||||||||||||
Db 1203717  GGCGGATAATCTTCTAAATCCATATAGCCGATCCAATAGATCTTCAGTCACGTTCAAGACAG  1203658

Qy    900  CCACCAAAATCCATATTTTGGGATCGGTTAATTAATTCAAGTGCAGGCGGTGCCAATATTACCGC  959
           |||||||||||||||||||
Db 1203657  TCGCTGCCGCGAGCTTTTAAGCTATTAAGTGTCTCAAGCTGAAAACTCTAGTA  1203598

Qy    960  CGACACCAACAATCATGCCTCCATGTTTTCGCCATCCTACTAATGTCGTAACAGTGC  1019
           |||||||||||||||||||
Db 1203597  CATAAAGTTCACAACTTCATTC---AACAATGACAAAGCGGGAAATCCCAATATATTCCGC  1203541

Qy   1020  TTTTGGCATTTGAGCCTGTGATGGCGGATAGATTGGTTGGTGTTGTTTGCTTGAGTTAGATCTC  1079
Db 1203540  CCATACCAACTTTCACACCAGCAGCTTTCGGCCATTCATAAACTAAAGTAGTTACGGTAC  1203481

Qy   1080  GTGCTTTGAGTGTATCAAATGAACAAATTGAACATCGCTTGATCACAGGAATGCCTTGAGCTTT  1139
Db 1203480  TTTTACCATTTGAACCTGTAATCCCCACAATTGGCTTTTGTCGCTGCTGCGGCA------  1203429

Qy   1140  TGGCGCGACAAATACTCGGCGGCGTTCTTGGGGTTTAATCGCCGGGCGCTGATGATCATCGG  1199
Db 1203428  -------------GAATAATTCAATATATCGCCGGATTACTTCCACTCCCGCTTTAA  1203388

Qy   1200  CTGAGCTTAAGAGTTCACTGTCCAAGCTGCCAAAATGACAAATGAACGCCTGTTGGGCCAGTT  1259
Db 1203387  GTGCGGGTTTGAAATTTCTGGTGTTTTTACCGCAAAGCCCTGGGCTTAATAACATCATATCGC  1203328

Qy   1260  TATCTGCCAAGGTAGGAGATTGGGATTACCATCGATGATGTCTAACTTTGTGACCGTGATTGA  1319
Db 1203327  TTTCAAGTAACCATTCCTGATTTAAACTACCAGTATGAAGAGGGATATTTTGAGGAAGTT  1203268

Qy   1320  CCAAAAATTCACACTGACAGTCCAGAGCTGCCCAAACCGACGACGGGCATTATTTTTT  1377
Db 1203267  TATCAATACAGTAGGATTTTTCGAGTATCAAATCACACGAATAATTAGCCTGTTGGGATA  1203208

Db 1203207  AGAGATAATCCACACAAAGAAAGACCTGTTTTGCCAAGCCCGATGATTAGTAATATATTTT  1203150
```

RESULT 11
US-10-981-687-1/c
; Sequence 1, Application US/10981687
; Publication No. US20050131222A1
; GENERAL INFORMATION:
; APPLICANT: Fleischmann et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Haemophilus Influenzae Rd Genome, Frag
; TITLE OF INVENTION: Thereof, and Uses Thereof
; FILE REFERENCE: PB186P2CID12
; CURRENT APPLICATION NUMBER: US/10/981,687
; CURRENT FILING DATE: 2004-11-05
; PRIOR APPLICATION NUMBER: US 10/158,865
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US 09/557,884
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 09/476,102
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: US 08/476,102
; PRIOR FILING DATE: 1995-06-07
; PRIOR APPLICATION NUMBER: US 08/426,787
;
```
;  PRIOR FILING DATE: 1995-04-21
;  NUMBER OF SEQ ID NOS: 1
;  SOFTWARE: Patentin version 3.1
;  SEQ ID NO 1
;  LENGTH: 1830121
;  TYPE: DNA
;  ORGANISM: Haemophilus influenzae
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (4747)..(4747)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (9921)..(9921)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (10150)..(10150)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (29298)..(29298)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (36543)..(36543)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (36551)..(36551)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (36636)..(36636)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (44905)..(44905)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (44975)..(44975)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (45593)..(45593)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (45732)..(45732)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (47036)..(47036)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (44416)..(44416)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (51334)..(51334)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
;  LOCATION: (51602)..(51602)
;  OTHER INFORMATION: n equals a, t, c, or g
;  FEATURE:
;  NAME/KEY: misc_feature
```

```
LOCATION: (51786)..(51786)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (51805)..(51805)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (55369)..(55369)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65309)..(65309)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (65313)..(65313)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (80024)..(80024)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (100091)..(100091)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (102696)..(102696)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (105121)..(105121)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (107248)..(107248)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (117136)..(117136)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119750)..(119750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (119924)..(119924)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (120038)..(120038)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121344)..(121344)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (121167)..(121167)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (122336)..(122336)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131340)..(131340)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (131360)..(131360)

OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (139910)..(139910)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (140398)..(140398)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (142750)..(142750)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145058)..(145058)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145171)..(145171)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (145942)..(145942)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (147197)..(147197)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (150841)..(150841)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
NAME/KEY: misc_feature
LOCATION: (152500)..(152500)
OTHER INFORMATION: n equals a,t,c, or g
FEATURE:
```

Query Match 11.6%; Score 160.4; DB 10; Length 1830121;
Best Local Similarity 49.6%; Pred. No. 1.4e-31;
Matches 683; Conservative 0; Mismatches 606; Indels 89; Gaps 7;

```
Qy        1 TAAAGCGTGAACCATGTTAACACAATTATCACCTCGCTCATTATAATTTTTAAATTGATC 60
              ||| || | || || |       |||  ||| |  || ||    | ||  ||   ||||
Db  1204439 TAAACATTGAGCTAAATGCCTAAAATTCTTCGCCCGGCTTTTCAAAAAGAAGACAAACTGATC 1204380

Qy       61 CATGCTAGCACAAGCAGGTGACAAAAGCACAGCCCCTAGCGGTGCTTTGACTGGCTGCCAA 120
              || | |||||||||||| | |||  ||| || || | | ||||   |||||  ||| |||
Db  1204379 GAGACTTGCACAAGCAGGCGCACAAGTAATACCA------TATCTCCGCTTTGCAA 1204332

Qy      121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGCCAAAGTGCATCACCTTT 180
              | |||| ||  |||     || || | |||  |  | | | ||||  | ||||||| ||
Db  1204331 TGTTGGGCGTAAAAAATTCTATCGCTTGTTCCATTGTATCGAACAA------ATAACTTTG 1204278

Qy      181 GAGTAAATCTTGCTCAAATTATCGGGGGCCATCTTGACCGATTAAATACACTATCCACCCA 240
              || | ||| || || ||||| | || |  |||||||||| |  |||| ||| ||| || ||
Db  1204277 CGATGAAAATTTGCAAGCAGCGCGCACCATCTGACCACCAAAAACAAATAATGTGTGG 1204218

Qy      241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACCTTTACCAAGCCACCCAAAAAT 300
              |||||| ||  | |||||    |||   |||||| |||| |||| || ||||||||||||
Db  1204217 TTGATTAATTAATTCAGCTAATTCTGAAAAATCAGCCCCTTTTCCGTCTCCGCCTAGCAA 1204158

Qy      301 CAAATGCCAGTGATGATGATATGCCATAAAACCGTGCCAAGCCCATCAATTGCTGCAATGGTTGA 360
              ||||||| | |    | ||||    ||||||   | |||||||| || |||| ||||| | ||
Db  1204157 CAAATGCAATTTACCCTCAATATA-------AAGCCCAGCCAAATGCAGCAACTGTACT 1204107

Qy      361 GCCGATGTTGTTGTACCTTTTGAATCATTAAAATAATATGCCTTCGCTATCAATATCATAATATA 420
              ||| ||||||||| ||||  | | |||||||   ||| || | |||| | ||||| |||| | |
Db  1204106 CCCCACACATTTGTTGCTTTAGAGTCATTAATCCAACGAAATGCCATTAGCTTGAATGCCGACTAA 1204047

Qy      421 TTCACAGCCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCAATGCCATCCATGGG 480
              ||| || | ||  | ||||| |||||||| | ||||||| ||||||| |||||||||||||
Db          
```

Query Match 11.2%;   Score 154.4;   DB 10;   Length 14324;
Best Local Similarity 49.0%;   Pred. No. 4e-31;
Matches 591;   Conservative 0;   Mismatches 546;   Indels 69;   Gaps 4;

Qy   1204046 TTGAAAACGATGATCTAACCCTTTGAAATGACGAAGTGCGGTACGAAGTGAATCTAAATT 1203987
Db   481 AAGATTGGCAGCTGCTCCTAAGGCAAGTGGCAGTAGGCAGTGAGTAGGGTTATGCTTGCC 540

Qy   1203986 AATACCTATAGCTTGTGTCGCCAATGCTGGTGCTGCCAAAAATGTTCATATATAATTATGGCGACC 1203927
Db   541 TTTTGATTTTTAATTGATGATGGTAAAGTTGTTGTTTTTTGATGATAAATGCCATACC 600

Qy   1203926 AACCAAATGTAGCTCTTCACAAGGTAAAATCACTTCATCTTTTACCATTAAAATATTGCTT 1203867
Db   601 ATCACCATCATCAAATAAATAAAAATCTGC-CGTTGGTGTGGCTAAGTAAGCTATTGTTG 659

Qy   1203866 GCCATTTCAGTTTTTAGCCAATAAATCCCGCCAAAAAATGTTCATATATAATTATGGCGACC 1203807
Db   660 TGATAAATGGTGTGATTGGGGGCAACGCTTCAGTGTGGTCAACGCATTGCTTGGCTAAAT 719

Qy   660 TGGGCATCGTCTTGGTGTAGAGATGGCGCGGGTTTGGCAATTTCAAAAATACCGCAATTTTGTG 779
Db   1203657 TCGCTGCGCCCGCAGCTTTTAAGCTATAAGTTGTCTCAAGCTGAAAAAAGCTCTAGTTA 1203598

Qy   780 CCACCAAAATCCCATATTTCGGATCGTTAATAATTCAAGTCGCAGGCGTGCCAAATTTACCCGC 959
Db   1203597 CATAAAGTTCACAAATCTCAATTC---AACAATGACAAAGGGGAATCCCAATATTTTTTGCC 1203541

Qy   960 CGACACCAACAATCATGCCTGCCGATCAAGATGATCAGCGGAGAGATTAAAATTG 1019
Db   1203540 CCATACCACTTTCACAACGTTTGCCCATTCATATAAAGTTAGTTACGGTAC 1203481

Qy   1020 TTTTGGCATTTGAGCTGCTGTGATGGGGATGAATTGGTGTGTTGTTGCTTGAGTTAGATCTC 1079
Db   1203480 TTTTTACCATTTGAATCCTGTAATCCCACAATTGGCTTTGTCTGCTGCGGTGCGGGGCA 1203429

Qy   1080 GTGCTTTGAGTGTATCAATGAACAAATTGAACATTGCCTGATCACAGGAATCGCCTTGAGCTT 1139
Db   1203428 ------------GAATAATTCAATATCGCGATTACTTCCACTCCCGCTTTAA 1203388

Qy   1140 TGGGCGGGTTTGAATTTCTGTGTTTTTACCGCAGCGCTTAATAACAATCATATCGC 1199
Db   1203387 GTGCGGTTTGAATTTCTGTGTTTTTACCGCAGCGCTTAATAACAATCATATCGC 1203328

Qy   1200 CTGAGCTTAAGAGTTCACTCTGCCAAAATGACAATGACGCCGTTGGCAATATACGCAA 1259
Db   1203267 TTTCAAGTAACCATTCCTGATTAAATACACCAGTACTGTATGAAGGGTAAAATGGTAA 1203208

Qy   1260 TAYCTGCCAAGGTAGGAATTGGGATTACCATCGATGATGTCGAACTTTGTGACCGTGATTGA 1319
Db   1203267 TATCAATACCAGTAGGATTTTTTCGAGTAATCAACTACCAGTAAAATGGTAA 1203208

Qy   1320 CCAAAAAATTCAGCTGACAGTCCCAAACCGACGACGGGGCCATATATTTGTT 1377
Db   1203207 AGAGATAATCCACACAAGAAAGACCTGTTTGTGCCAAGGCCGATGATAGTAATATTTT 1203150

Qy   172 ATCACCTTTGAGTAAATCTTGCTCAATTATCGGGGGCCATTCTTGACCGGATTAAATACACT 231
Db   5280 ATAGCTTTGTGATGAAAATTTTGCAAGCTGCACACCAATTACCGACCAAAAAAACAAAT 5339

Qy   232 ATCCACCCACTGCCATAACATATGTGATTAATTAAATTACTTGACACACTTTAAGGCCC 291
Db   5340 AATGTGTGGTTGATTAATTAATTCAGCTAATTCGAAAAATCAGCCCCTTTTCCGTCTCC 5399

Qy   292 ACCCAAAAATCAAATGCCAGTGATGATGATAAGTGCCATATGGCAAGCCCATCAAATTGCTGC 351
Db   5400 GCCTAGCACAACAAAATGCAAT-------TTACCCTCAATTATAAAAACGTGCACGTAATGCAGC 5450

Qy   352 AATGGTTGAGCCGATGTTTGTGTACCTTTTGAATCATTAAAAATATCGGCTTGCTATCAATATC 411
Db   5451 AACTGTACTCCCCACCAATTGTTGCTTTAGAGTCATTAATCCAACGAATACCAATTGCTTG 5510

Qy   412 ATCAATATATTCACAGCCGATGCGGTAGCCCGTTTGCCATCTAAAATCTTGAAAGTCATAGGC 471
Db   5511 ATGCACTAATTGAAAAACGATGAATCAAACCCTTTGAAAGTGCGGTATCAAGCATGGA 5570

Qy   472 ATCCATGGGAAGAGATTGGCAGCGTGCCTAAGGCAAGAGTGCGAGATAAGGCATTGAGTAGGTT 531
Db   5571 ATCTAAATTAATACCGACACGCTTGTGTCGCCAATAATCGGCTATATTTTGCTATATAATT 5630

Qy   532 ATGCTTGCCTTTGATTTTTAATTGATGATGGTTTTTGTTTTTTGATGATAA 591
Db   5631 ATGGGCGACCGACTAATGTTCACTTCTTCACAAGGCAAAATTTACTTCATCTTTTTTGATGATAA 5690

Qy   592 TGCCATACCATCACCATCATCAAATAAATAAAAATCTGCCAATAAATCGGCTAGAAGCT 651
Db   5691 GTATTGCTTGGCCATTTTCAGTTTTTAGCCAATAAATCGGCTAGAAAAGTTGAAGC 5750

Qy   652 ATTTGTTGTTGATAAATGGTGGATTGGGGGCAACGGCTTGTCAGTGTGGTCAAGCATTGCTT 711
Db   5751 GGTTGTCTCGGCTTGATTTCACCTTCACCTTAAAAAAGTCAG------- 5790

Qy   712 GGCTAAAGTTGCATCGTCTTGGTAGATGATGGGCGTTGGCAATTTTCAAAATACGCCAA 771
Db   5791 ------TTATCCCTCATTATTCAAAAACTCCGACGACCTCAGCATTGAATAAAATGGTAA 5840

Qy   772 TTTTTGTGCCAAATAATCACCATGCCATCGTGCGATCAAGAGATCACGCGGAGAGATT 831
Db   5841 TTTTGCTTGGCGGATAATCTTCTAAAATCCATATAGGCCATAATCTTCAGTCACGTT 5900

Qy   832 TAAAATTGTTGCCCCTTGAGCTGCCAAATTTGAGAGTATGATATAGTCAAGCTGAAAACTTGACAG 891
Db   5901 CAAAAACAGTTGCAGCGCCGATGCGTTTAAGCTATAAGTGTCAAGCTGAAAAACTTTAGAAATG 5960

Qy   952 ATTACCGCCGCGACACCAAATCATCAATATTTTGGATCGTTAATAACTCAAGTTCGCATCTGG 1011
Db   5961 CTCTAGTACATAAAAGTTCACAAATCTTCATTC---AACAATGACAAAAGCGGGGAAATCCAAT 6017

Qy   1012 AACAGTGCTTTGGCATTTGAGCCTGTGATGGGGCCGATGATGGGTCGTGTTGTTGCTTGAGT 1071
Db   6018 ATTTCCGCCCCATACCAACTTTTCACGCGGCCATTTCGCCCATTTCAAAAACTAAGTAGT 6077

Qy   1012 AACAGTACTTTTTCCATTTGAACCTGTAATCCCCCAAATTTGGCTTTGTCTGCTTGTTGCGCG 1071
Db   6078 AACAGTACTTTTTCCATTTGAACCTGTAATCCCCCAAATTTGGCTTTGTCTGCTGCGCG 6134

RESULT 13
US-10-795-159-685/c
Sequence 685, Application US/10795159
Publication No. US20050221439A1
GENERAL INFORMATION:
APPLICANT: BAKALETZ et al.
TITLE OF INVENTION: GENES OF AN OTITIS MEDIA ISOLATE OF NONTYPEABLE H. INFLUENZAE
FILE REFERENCE: 2835/38815A
CURRENT APPLICATION NUMBER: US/10/795,159
CURRENT FILING DATE: 2004-03-05
PRIOR APPLICATION NUMBER: US 60/453,134
PRIOR FILING DATE: 2003-03-06
NUMBER OF SEQ ID NOS: 771
SOFTWARE: Patentin version 3.2
SEQ ID NO 685
LENGTH: 908766
TYPE: DNA
ORGANISM: H. influenzae
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9223)..(9223)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (39640)..(39640)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (292404)..(292404)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (292495)..(292495)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (908604)..(908641)
OTHER INFORMATION: n is a, c, g, or t
US-10-795-159-685

Query Match 11.2%; Score 154.4; DB 10; Length 908766;
Best Local Similarity 49.0%; Pred. No. 4.4e-30;
Matches 591; Conservative 0; Mismatches 546; Indels 69; Gaps 4;

Qy   172 ATCACCTTTGAGTAAATCTTGCTCAATTATCGGGCATCTTGACCGATTAAATACACACT 231
           ||| |||| ||| || || || ||| | | |||| ||| ||||||| | |||
Db 107917 ATAGCTTTGTGATGAAAATTTTGCAAGCTGCACCACCATTCTCGACCAAAAACAATACAAAT 107858

[Alignment blocks continue — query (Qy) and database (Db) nucleotide sequences with match bars, positions 232/107857 through 1372/106847.]

```
Qy  1312  GTGATTGACAAAAAATTCAGCTTGACAGTTCCAGAGCGTGCCAAAACCGACGACGGCGCATA  1371
                ||||  |||||  || |||  |||    |||||
Db 106846 TTGGGATAGAGAGATATCCACACAAGAAAGACCTGTTTTGCCAAGCCAATGATCGTAAT  106787

Qy  1372  TTTTTTT  1377
          |||||
Db 106786 ATTTTT  106781
```

RESULT 14
US-10-282-122A-25165//c
Sequence 25165, Application US/10282122A
Publication No. US20040029129A1
GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; SEQ ID NO 25165
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 25165
; LENGTH: 1281
; TYPE: DNA
; ORGANISM: Legionella pneumophila
US-10-282-122A-25165

Query Match 11.1%; Score 152.8; DB 8; Length 1281;
Best Local Similarity 51.6%; Pred. No. 2.8e-31;
Matches 464; Conservative 0; Mismatches 412; Indels 24; Gaps 4;

Qy 170 GCATCACCTTTGAGTAAAACTTGTCTCAATATATCGGGGGCATCTTGACCGATTAAATACACA 229
Db 1163 GGAACTACTTTGGCTAAAGCGGATTCAAATCTTGTTCAGCATCTTCACCAATCAAAACAATG 1104
Qy 230 CTATCCACCACCTGCCATAACATATGCTGACAGCTTCACTAAAAATCTAAAAATCTTGACCTTTTACCAAGC 289
Db 1103 GAACGAACAAAATTCTGAAAACTGGCTGGGCTAATTCTTTGGAAAATCCGCCCTTTTCCCTGT 1044
Qy 290 CCACCCAAAATCAATGCCAGTGATGATATATCCATAAAACCGTGCCAAGCCCCATCAATTGCT 349
```

```
Db 1043  CCACCCGCAATAAGAACAACAATTTCCCTTGCATAGAACC---TCCTATTCCATTAATAGCA  987
Qy  350  GCAATGGTTGAGCCGATGTTGTGTACCTTTGAATCATTAAAATATGCCTTGCTATCCAATA  409
Db  986  GAGATTGTAGCACCAATATTTGTTCCTTTGAAATCATTAATCAACCTACTCCATCTACT  927
Qy  410  TCATCAATATATTCAACAGCGATGCGGTAGCCCCTTTGAAAGTTCTTAAGGGGTATCAAGCATG  469
Db  866  TGTTGCATAGAAAATTCCGCTGCTTCGGCTAAAGCACAAGCAGTGCGAGATAAGGCATTGAGTAGG  807
Qy  470  GCATCCATGGGAAGATTGGCAGCCTTGTCCTAAGGCAAGTGCCGAGATAAGGCATTGAGTAGG  529
Db  806  TTATGCACCCCTTTATTAAAATCGATTCCACTGGGCAAAAGAGCGCTCCATTCCTTTCGCC  747
Qy  530  TTATGCTTTGCCTTTGAGTTTTGAATTCATCGGATTGGTTAAAAGTTTGTTTTTTGATGATAT  589
Db  746  AAATAAGTAGTATTCTCTTGTTCAATCAGCCCCAATTTCCCATT-------GAAG  698
Qy  590  AATGCGATACCATCACCATCATCAAATATAAATTCTCGCGTTTGTTGGTGCGTTAAG  649
Db  584  TGTTTTTGCTTCGGTATAAGGCTTCCATAGTGTGATAGCTTGTCCAAAATGATCAGGGCGTTACA  525
Qy  650  CTATTGTTGTTGATAATGGTGTGATTGGGGCAACGGCTTGTCAGTGTGTGGTCAAGCAGTTGC  709
Db  641  ---ACAGTGTAGACGTCTTCACGATTAAATAAAACAGCCCTGGCACACGATAAAATCCTT  585
Qy  710  TTGGCTAAATTGGCATCGTCTTGGTAGATGATGGCCGGTTGGCAATTTCAAAAATACGC  769
Db  697  GCGCGTTCTTTGGCCAAAAGAGATGCAATTTTATATCGTGTTGACAGGATTGGTGTGAGGA----  642
Qy  770  AATTTTTGTGCCAAAATAAATCCACACAATGCCATCGTGCCGATCAAGATGATCAGGGCGAGAGA  829
Db  524  TTAAGAAATCGTCGCAACCACTGGCGACAAGAATAAGTTAAATCCAAATTGAAAAACTTGGAT  465
Qy  830  TTTAAAATTGTTGCCCCTTGAGCTGCCAAATTTGAGATATGCTCAAGCTGAAAAACTTGAC  889
Db  464  AACTCCAATAACCACAAATCGTAGTGAATGTTCATCATCATATCTAAGACAGGAGTG  405
Qy  890  AGCTCAAGCACCACCAAAATC---CATATTTTGGATGTGTTAATAATTCAAGTGCAGCGTG  946
Db  404  CCGATATTTCCGGCTACAGCTACGGCGAAACCCGGCGCCTTAGCCATTTCTCCCACCAAA  345
Qy  947  CCAAATATTACCGCGCAACCGACACACAACATCATCGCCTGCATGTTTTGCCCTACTAAT  1006
Db  344  GTCGTTACAGTGGATTACCATTAGTTCCCGTTTATCGCAATGACAGGGGGCGTAATCTCT  285
Qy 1007  GTCGTAACAGTGCTTTTGGCATTGAGCCTGCTGATGAGTTGGGTGTGTTTGTTGCT  1066
```

RESULT 15
US-10-958-216-402//c
Sequence 402, Application US/10958216
Publication No. US20050181388A1
GENERAL INFORMATION:
; APPLICANT: EDWARDS, ALED
; APPLICANT: DHARAMSI, AKIL
; APPLICANT: VEDADI, MASOUD
; APPLICANT: ALAM, MUHAMMAD ZAHOOR
; APPLICANT: ARROWSMITH, CHERYL
; APPLICANT: AWREY, DONALD E.
; APPLICANT: BEATTIE, BRYAN
; APPLICANT: BUZADZIJA, KRISTINA
; APPLICANT: CANADIEN, VERONICA
; APPLICANT: DOMGALA, MEGAN
; APPLICANT: HOUSTON, SIMON
; APPLICANT: KANAGARAJAH, DHUSHY
; APPLICANT: LI, QIN
; APPLICANT: MANSOURY, KAMRAN
; APPLICANT: MCDONALD, MERRY-LYNN
; APPLICANT: NETHERY, KATHLEEN
; APPLICANT: NG, IVY

```
; APPLICANT: OUYANG, HUI
; APPLICANT: PINDER, BENJAMIN
; APPLICANT: RICHARDS, DAWN
; APPLICANT: TAI, MATTHEW
; APPLICANT: THALAKADA, ROSANNE
; APPLICANT: VALLEE, FRANCOIS
; APPLICANT: VIRAG, CRISTINA
; TITLE OF INVENTION: NOVEL PURIFIED POLYPEPTIDES FROM BACTERIA
; FILE REFERENCE: IPT-205.01
; CURRENT APPLICATION NUMBER: US/10/958,216
; CURRENT FILING DATE: 2004-10-04
; PRIOR APPLICATION NUMBER: PCT/CA03/00462
; PRIOR FILING DATE: 2003-04-02
; PRIOR APPLICATION NUMBER: 60/369,511
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 60/385,089
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/385,751
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/385,553
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/385,577
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/386,367
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/386,566
; PRIOR FILING DATE: 2002-06-05
; PRIOR APPLICATION NUMBER: 60/386,390
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: 60/386,601
; PRIOR FILING DATE: 2002-06-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1132
; SOFTWARE: Patentin Ver. 3.3

SEQ ID NO 402
  LENGTH: 1314
  TYPE: DNA
  ORGANISM: Haemophilus influenzae
US-10-958-216-402

Query Match      11.0%;  Score 152.4;  DB 10;  Length 1314;
Best Local Similarity   49.2%;  Pred. No. 3.7e-31;
Matches  678;  Conservative  0;  Mismatches 611;  Indels  89;  Gaps  7;
```

```
Qy       1 TAAAGCGTGAACCATGTTAACAAACTTATCACCTCGCTCATTATATAATTTTAAATTGATC 60
             ||||      |||||||      ||    |||| |  |||||||||||     |||||||
Db    1308 TAAACATTGAGCTAAAATGGCTAAAATTCTTCGCGCCGCTTTTCAAAGAGAGCAAACTGATC 1249

Qy      61 CATCCTTAGCACAAGCAGGTGACAAAAGCACAGCCCTAGCGGTGCTTTGACTGGTGCCAA 120
             |   ||||||||||||||| || || ||||||||||  |  || ||| || |  |||||
Db    1248 GAGACTTGCACAAGCAGGGCGGACAATAATACCA------TATCTCCGCTTTGCAA 1201

Qy     121 TACTGCCCGATCAAGTGTACCAACCTGATATATTTTAGATGTCGCCAAATGTGCATCACCTTT 180
            ||| | || |  ||   | |  |||| | | || ||| |||  |||||  |
Db    1200 TGTTTGGGCCGTTAAAAAATTCTATACGCTTGTTCCATTGTTTATCGAACCAA------ATAACTTTTG 1147

Qy     181 GAGTAAAATCTTGCTCAATTATTCGGGGGCCATCTTGACCGGATTAAATACACTACTTCCACCCA 240
            | ||||||  |||  ||| | |   |||| |||    |||  || ||||  || |  |||||| |
Db    1146 CGATGAAAAATTTGGAAGCAGCGCACCCATCTCGACCAAAAACAATTAACAAATAATGTGTGG 1087

Qy     241 CTGCATAACATATGCTGACAGCTCACTAAAATCTTGACTTTACCAAGCCCACCCAAAAT 300
            |||||||| ||||| || || ||| ||| ||||  |||||| | |||| ||| |||||
Db    1086 TTGATTAATTAATTCGAAAAAATCAGCCCCTTTTCCGTCTCGCCCTAGCAA 1027

Qy     301 CAATGCCAGTGATGATATGCCATAAAACCGTGCCAAGCCCATCAATTGCTTGCAATGGTTGA 360
            |||| ||  | ||   ||  || |||| |||  ||| |||| ||||  |  || ||  | |||
Db    1026 CAAATGCAATTACCCTCAATATA-------AGCCCAGCCAATGCGCAGCCAACTGTACT 976

Qy     361 GCCGATGTTTGTACCTTTTGAATCATTAAAAATATGCCTTGCTACTAATATCATCAATATA 420
            |||  ||||||| || || ||||||||||   ||||| |||| ||| |||| ||||| ||||
Db     975 CCCCACATTTGTTGCTTAGAGTCATTAATCCAACGAAATGCCATTAGCTTGAATGCACTAA 916

Qy     421 TTCACAGCGATGCGGTAGCCCTTTGAAAGTCTTAAGGGTATCAAGCAATGGCATCCATGGG 480
            |||||||| |||||||||||| ||||||| ||||| || |||||||| ||||| | |||||
Db     915 TTGAAAACGGATGAATCTAACCCTTTGAAAATGACGAAGGTGCGGTACGGAATTGGAATCTAAAATT 856

Qy     481 AAGATTGGCAGCTGTGCCTAAGGCAAGTTGCAGAATGTGAGAATAAGGCATTGAGTAGGTTATGCTTGCC 540
            |||||||||||| |||||||||||||||||| |||| |||||||||| ||||| ||||| || ||||||||
Db     855 AATACCTATAGCTTGTGCCAATGCTGTTGCTGCCAAAAATGTTCATATAATTATATGGGAACC 796

Qy     541 TTTGATTTTTAATTGATGGATTGGTAAAAGTTGTTTTTTTGATGATATAATGCCATACC 600
            |||||||| |||||||||| ||||| ||| ||||  |||||  | |||| |||| || |||
Db     795 AACCAATGTAGCTTCTTCACAAGGTAAAATCACTTCATCTTTTACCATTAAAATATTGCTT 736

Qy     601 ATCACCCATCATCAAATAAATAAAAATCTGC-CGTTTGGTGGCTAGAGTAAGCTATTTGTTG 659
            ||||||||||||||||||||||||  |||||  | ||| |  |||  | ||||| ||||  ||
Db     735 GCCATTTTCAGTTGTTTTTAGCCAATAATCCGACTATTTTCCGCAAAAGAAAACGGGTATGTTT 676

Qy     660 TGATAAATGGTGTGATTTGGGGCACAGGCTTGTCAGTGTGGTCAAGCATTGCTTGGCTAAAT 719
            ||    |||| | ||||| |||||||||||||||||||||||||||||||| |||||| |||
Db     675 CGCTTGATTTTCGTTGTTTTCCCAAAAGTCAGGC------ 644

Qy     720 TGGCATCGTCTTGGTAGAATGAATGGGCGGTTTGGCAATTTTCAAAAATACGCAATTTTTGTG 779
            |||||||||| |||||||||||||||||| || |||||||||||||| |||||  |||  |||
Db     643 ---TATCTTCATTGTTCAACACACTACTTTAGCATTATGACATTAAAATGCGTAAATTTTGCTT 587

Qy     780 CCAAATAATCCACCATCCATGCGTGGCCGATCAAGATGATCAGCGGAGAGATTTAAAAATG 839
            || ||||||||  || ||||| ||| |||||||||||||||||| |||||  || ||||||
Db     586 GGCGGATAAATCTTCTCAAATCCATATAGCGGATCCAATATTTCAGTCAGTCACGTTCAAGACAG 527

Qy     840 TTGCCCCTTGAGCTGCCAAATTTGAGAATGCTCAAGCTCAAAAAACTTGACAGCTCAAGCA 899
            ||||| |  | ||| ||||||||||||||||||||| ||||||||| ||||| |||||||||
Db     526 TCGCTCGCCGCAGCTTTTAAGGCTATAAGCGTGTCTCAAGGCTGAAAAACTCTAGTA 467

Qy     900 CCACCAAAATCCATATATTTTGGATCGTTTAATTAATTCAAGTGCAGGCGTGCCAATATTATACCGC 959
            ||| |||||||||||| |||| ||| |||||| ||| |||||  || ||||| |||||||||||||
Db     466 CATAAAGTTCACAATACCCTCATTC---AGCAATGACAAAAGCGGGGAATCCCAATATTTCCGC 410

Qy     960 CGACACCAACAATCACGCCTGCATGTTTTGCCATCCTGCGTTTGTGTTTTGCTTGAGTTAGATCTC 1019
            | | ||||| | | ||||  || |||||||||||  || ||||| ||||||| |||| | |||||
Db     409 CCATACCAACTCTACTGTAGCACGTTTCGCCATTTCATAAAACTAAAGTCGTTACGGTAC 350

Qy    1020 TTTTTTGCCATTTGAGCCTGTGAGTGGGGATGAGATTGGGTGTGTGTTTGTGCTTGAGTTAGATCTC 1079
            ||||| ||||||| |||||| |  ||  ||| || ||||| | |||   | | ||   | |||| |
Db     349 TTTTTACCATTAGATCCCTGTAATCCCCACAATTGGCTTTGTGCGCCTGCGGCCGAC-------- 298

Qy    1080 GTGCTTTGAGTGTATCAATGAACAATTGAACATCGCTGATCACAGGAAATGCCTTGAGCTT 1139
                        ||||| |||||||||||| |||| | || || |||| |||| |  |||||
Db     297 -----------GAATAATTCAATAATCGCCGATTACTTCCACTCCGCGTTAA 257

Qy    1140 TGGCGCGGCGACAATACTCGGCGGTTCTTCGGGGTTAAATGCGCGGGGCTGATGATCGATTCGATCGG 1199
            ||| |||| ||||||||| |||||| ||| |   || ||||||||||| |||||  ||  | |||||
Db     256 GTGCGGTTTGAATTTCTGGTGTTTTTTTACCGGCGAAGCCCTGGGCTAATAAACAATCATATCGC 197

Qy    1200 CTGAGCTTAAGAGTTCACTGTGCCAAAATGACAAATGAACGCCTGTGGGCGGAGTT 1259
            || | ||||| | |||| |  | ||||||||||||||| |   || ||| || |||
Db     196 TTTCAAGTAACCATTCCTGATTTAAACTACCAGTATGATGAAGGGGATATGTTTGTGACGGTGGATA 137

Qy    1260 TATCTGCCAAGGTAGGATTGGGATTACCATCGATGATGCTAACTTTGTGACGTGATTGA 1319
            ||||  |||| | || ||| | |||||||||||||| |||||||||||||| ||| |||
Db     136 TATCAATACCAGCAGGAGATTTTTTCGAGTATCAATCACACGAATATTAGCCTTGTGTGGGATA 77

Qy    1320 CCAAAAAATTCACACGCTGACAGTCCCAAAACCGACGACGGGCCATAATATTTTTT 1377
            |||| ||||||||||  | ||||  |||||||| | ||||| || ||||| ||||||
Db      76 AGAGATAATCCACACAAGAAAGAACCTGTTTTGCCCAGCCCATGAGTAGTATATTTTT 19
```

10-672-787-35_11357_12736.minsz22.rnpbn

GenCore version 5.1.8
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 18, 2006, 09:34:46 ; Search time 51 Seconds
                 (without alignments)
                 1909.160 Million cell updates/sec

Title: US-10-672-787-35_COPY_11357_12736
Perfect score: 1380
Sequence: 1 taaagcgtgaaccatgttaa...........acgacggcatattttttga 1380

Scoring table: IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0
Searched: 224314 seqs, 35277956 residues

Total number of hits satisfying chosen parameters: 433684

Minimum DB seq length: 22
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications NA_New:*
                 1: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US09_NEW_PUB.seq:*
                 2: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US06_NEW_PUB.seq:*
                 3: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US07_NEW_PUB.seq:*
                 4: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US08_NEW_PUB.seq:*
                 5: /EMC_Celerra_SIDS3/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
                 6: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US10_NEW_PUB.seq:*
                 7: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US11_NEW_PUB.seq:*
                 8: /EMC_Celerra_SIDS3/ptodata/1/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34.2	2.5	1302	7	US-11-217-529-118	Sequence 118, App
2	32.6	2.4	474	7	US-11-217-529-2634	Sequence 2634, Ap
3	32.2	2.3	3411	6	US-10-504-120-2	Sequence 2, Appli
4	31.8	2.3	516	7	US-11-270-287-3	Sequence 3, Appli
5	31.6	2.3	1632	7	US-11-217-529-77328	Sequence 77328, A
6	31.6	2.3	789	7	US-11-217-529-3777	Sequence 3777, Ap
7	31.6	2.3	1125	7	US-11-217-529-3776	Sequence 3776, Ap
8	31.4	2.3	3653	6	US-10-504-120-16	Sequence 16, Appl
9	31.4	2.3	5026	7	US-11-217-529-166180	Sequence 166180,
10	31.2	2.3	1176	7	US-11-217-529-1381	Sequence 1381, Ap
11	31.2	2.3	1629	7	US-11-217-529-76422	Sequence 76422, A
12	31.2	2.3	2385	7	US-11-217-529-4571	Sequence 4571, Ap
13	31.2	2.3	1077	7	US-11-217-529-80459	Sequence 80459, A
14	31	2.2	1593	7	US-11-217-529-78136	Sequence 78136, A
15	30.8	2.2	1125	7	US-11-217-529-79486	Sequence 79486, A
16	30.8	2.2	1722	7	US-11-217-529-3849	Sequence 3849, Ap
17	30.6	2.2	1311	7	US-11-217-529-720	Sequence 720, App
18	30.4	2.2	455	7	US-11-301-554-720	Sequence 720, App
19	30.4	2.2	459	7	US-11-301-554-625	Sequence 625, App
20	30.4	2.2	1353	7	US-11-217-529-79251	Sequence 79251, A
21	30.4	2.2	1896	7	US-11-217-529-75758	Sequence 75758, A
22	30.2	2.2	3540	7	US-11-217-529-2192	Sequence 2192, A
23	30.2	2.2	1692	7	US-11-217-529-80706	Sequence 80706, A
24	30.2	2.2	4899	7	US-11-217-529-4664	Sequence 4664, Ap
25	30.2	2.2	15456	7	US-11-324-284-23	Sequence 23, Appl
26	30	2.2	1665	7	US-11-217-529-3737	Sequence 3737, Ap
27	30	2.2	1764	7	US-11-217-529-79615	Sequence 79615, A
28	29.8	2.2	5778	7	US-11-217-529-5615	Sequence 5615, Ap
29	29.8	2.2	565	7	US-10-488-619-1129	Sequence 1129, Ap
30	29.8	2.2	567	6	US-10-488-619-2704	Sequence 2704, Ap
31	29.8	2.2	578	6	US-10-488-619-1398	Sequence 1398, Ap
32	29.8	2.2	582	6	US-10-488-619-1435	Sequence 1435, Ap
33	29.8	2.2	591	6	US-10-488-619-1401	Sequence 1401, Ap
34	29.8	2.2	656	6	US-10-488-619-1128	Sequence 1128, Ap
35	29.8	2.2	774	6	US-10-488-619-1434	Sequence 1434, Ap
36	29.6	2.1	1000	7	US-11-233-726-4	Sequence 4, Appli
37	29.6	2.1	1182	6	US-10-505-928-442	Sequence 442, App
38	29.6	2.1	1182	6	US-10-511-937-2869	Sequence 2869, Ap
39	29.6	2.1	1353	7	US-11-217-529-2417	Sequence 2417, Ap
40	29.6	2.1	2008	7	US-11-246-999-26	Sequence 26, Appl
41	29.6	2.1	2682	7	US-11-217-529-432	Sequence 432, App
42	29.6	2.1	3972	6	US-10-473-173-131	Sequence 131, App
43	29.4	2.1	2196	7	US-11-217-529-1751	Sequence 1751, Ap
44	29.4	2.1	4221	7	US-11-217-529-2728	Sequence 2728, Ap
45	29.4	2.1	128361	6	US-10-505-928-151	Sequence 151, App

ALIGNMENTS

RESULT 1
US-11-217-529-118/c
; Sequence 118, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 118
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-118

  Query Match   2.5%; Score 34.2; DB 7; Length 1302;
  Best Local Similarity   50.3%; Pred. No. 0.42;
  Matches   84; Conservative   0; Mismatches   83; Indels   0; Gaps   0;

Qy   446 AAAGTCTTAAGGGTATCAACGACATGCCATCGGGAAAGATTGGCAGCTGTGCCTAAGGCA 505
              ||||||| | ||| ||| | | ||| | | | | || || ||| | | ||||
Db   1147 AAAGGTGTAATGGTTATAAAACAGAGGCTATACCGGTTTAAAATTTGGACGTGACGAAAAGAC 1088

Qy   506 AGTGCAGATAAGGCATTGAGTAGGTTATGCTTGCCTTTGATTTTTAATTGATGGATTGGT 565
              ||| | | || || |||| || | || | | | ||| | ||||| | ||| |
Db   1087 ATACCATGATAACCAAGGATGTCTGGTAATCCATACTAGGGGCTTCAAAATTGTAGGATCGGGGT 1028

Qy   566 AAAAGTTTGTTTTTTGATGATATAATGCCATACCATCACCATCATC 612
              | | ||| |||||| || ||| | | | || || || |
Db   1027 AGATTTTGCTAAAATATATGGAGGAATATAGTCAAATTAATAATATATTC 981

RESULT 2
US-11-217-529-2634/c
; Sequence 2634, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED

```
;    APPLICANT: NAKAO, YOSHIHIRO
;    APPLICANT: NAKAMURA, NORIHISA
;    APPLICANT: KODAMA, YUKIKO
;    APPLICANT: FUJIMURA, TOMOKO
;    APPLICANT: ASHIKARI, TOSHIHIKO
;    TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;    FILE REFERENCE: S-38-285
;    CURRENT APPLICATION NUMBER: US/11/217,529
;    CURRENT FILING DATE: 2005-09-02
;    PRIOR APPLICATION NUMBER: US 10/932,182
;    PRIOR FILING DATE: 2004-09-02
;    NUMBER OF SEQ ID NOS: 197023
;    SOFTWARE: Patentin version 3.3
;
;    SEQ ID NO 2634
;    LENGTH: 474
;    TYPE: DNA
;    ORGANISM: Saccharomyces pastorianus
     US-11-217-529-2634

     Query Match             51.0%;  Score 32.6;  DB 7;  Length 474;
     Best Local Similarity   51.0%;
     Matches   77;  Conservative    2.4%;  Mismatches  0.73;  Pred. No. 74;  Indels  0;  Gaps  0;

Qy    907  ATCCATATATTTGGATCGTTAATAATTCAAGTGACGTGCCAATATATTACCGCCGACACC  966
                |||  ||   ||||  | ||| ||||| |||  || |  || ||| |  |  | |  | ||
Db    459  ATCAAAAGTTTGGCCTTCGACTAATTCAGGAATAGCTTCGTCCTTCTTTTCAGCGTCAGC  400

Qy    967  AACAATCATGCCTGCATGTTTTGCCATCTCGCCTACTACTAATGTCGTAACACAGTGCTTTTGGC  1026
                | |||| || || |  ||  | || ||| ||||  |  |||||||||| |||  |||| ||| ||
Db    399  TGGAGCATTGGCTTCGTCGCTTTTTCCATTTGGGCGTCCTGCCCAATTGGGACAAGGCCTGGATGGC  340

Qy   1027  ATTTGACAGCCTGTGATGGCGATGATTCGGTTGGTGTG  1057
                |||  |  ||| | |||   |  | |||| | |||
Db    339  TTCAGGGCCCAATTCGGAGATAATCCTGGG  309


RESULT 3
US-10-504-120-2/c
Sequence 2, Application US/10504120
Publication No. US20060088829A1
GENERAL INFORMATION:
;    APPLICANT: Exelixis, Inc.
;    TITLE OF INVENTION: MINRs AS MODIFIERS OF INSULIN RECEPTOR SIGNALING AND METHODS OF
;    USE
;    FILE REFERENCE: EX03-003C-PC
;    CURRENT APPLICATION NUMBER: US/10/504,120
;    CURRENT FILING DATE: 2004-08-06
;    PRIOR APPLICATION NUMBER: 60/354,824
;    PRIOR FILING DATE: 2002-02-06
;    PRIOR APPLICATION NUMBER: 60/358,217
;    PRIOR FILING DATE: 2002-02-20
;    PRIOR APPLICATION NUMBER: 60/358,189
;    PRIOR FILING DATE: 2002-02-20
;    PRIOR APPLICATION NUMBER: 60/358,126
;    PRIOR FILING DATE: 2002-02-20
;    PRIOR APPLICATION NUMBER: 60/358,995
;    PRIOR FILING DATE: 2002-02-21
;    PRIOR APPLICATION NUMBER: 60/358,765
;    PRIOR FILING DATE: 2002-02-21
;    PRIOR APPLICATION NUMBER: 60/358,756
;    PRIOR FILING DATE: 2002-02-21
;    PRIOR APPLICATION NUMBER: 60/359,531
;    PRIOR FILING DATE: 2002-02-25
;    PRIOR APPLICATION NUMBER: 60/360,222
;    PRIOR FILING DATE: 2002-02-26
;    PRIOR APPLICATION NUMBER: 60/360,222
;    PRIOR FILING DATE: 2002-02-26
;    PRIOR APPLICATION NUMBER: 60/360,224
;    PRIOR FILING DATE: 2002-02-26
;    Remaining Prior Application data removed - See File Wrapper or PALM.
;    NUMBER OF SEQ ID NOS: 32
;    SOFTWARE: Patentin version 3.2
;    SEQ ID NO 2
;    LENGTH: 3411
;    TYPE: DNA
;    ORGANISM: Homo sapiens
     US-10-504-120-2

     Query Match             53.2%;  Score 32.2;  DB 6;  Length 3411;
     Best Local Similarity   53.2%;
     Matches   92;  Conservative    2.3%;  Pred. No. 2.9;  Mismatches  78;  Indels  3;  Gaps  1;

Qy    136  TGTACCAACCTGGATAATATTTTAGATGCCAA--ATGTGCATCACCCTTTGAGTAGAATCTTG  192
              || | || | | |||  |   || | ||  || |   ||| | || |||| || |||  |||
Db   1335  TTTATCACCATGACCTTATTCGAGAGTGCAAAATTTATGTCCAAAAACTTATCCATTCTTTTG  1276

Qy    193  CTTCAATTATCGGGCGCATTCTTGACCGGATTAAATACACACTTATCCACCCACTGCATAACATA  252
              ||  | ||| || | |  ||  || | | ||||||| ||| | || || | || |  || ||
Db   1275  TACCAGTATTCGAACCCCTTCAAATGCTCCTTATAGAAGCTAATCCACATCAGCAATGGCCAA  1216

Qy    253  TGCTGACAGCTCACTAAAAATCTTGACCTTTACCAAGCCCACCCAAAAATCAATG  305
              ||| || ||| | ||   |  || |  ||  || | |||  || ||| ||  |
Db   1215  GGATGTCAGCTGAGCAGTCCCTGTCCCATCCGTCACTGCAATGACAAGCACTG  1163


RESULT 4
US-11-270-287-3
Sequence 3, Application US/11270287
Publication No. US20060099627A1
GENERAL INFORMATION:
;    APPLICANT: Kara, Anna K.
;    APPLICANT: Ting, Robert C.
;    APPLICANT: Tham, Jill M.
;    APPLICANT: Nelson, James S.
;    APPLICANT: Tan, Theresa M.
;    TITLE OF INVENTION: Diagnosis of Parasites
;    FILE REFERENCE: 64-99
;    CURRENT APPLICATION NUMBER: US/11/270,287
;    CURRENT FILING DATE: 2005-11-09
;    PRIOR APPLICATION NUMBER: US/09/369,992
;    PRIOR FILING DATE: 1999-08-06
;    PRIOR APPLICATION NUMBER: PCT/IB98/00212
;    PRIOR FILING DATE: 1998-02-05
;    PRIOR APPLICATION NUMBER: AU PO9481/97
;    PRIOR FILING DATE: 1997-09-26
;    PRIOR APPLICATION NUMBER: AU PO9329/97
;    PRIOR FILING DATE: 1997-04-21
;    PRIOR APPLICATION NUMBER: AU PO4953/97
;    PRIOR FILING DATE: 1997-02-06
;    NUMBER OF SEQ ID NOS: 53
;    SOFTWARE: Patentin Ver. 2.0
;    SEQ ID NO 3
;    LENGTH: 516
;    TYPE: DNA
;    ORGANISM: Plasmodium berghei
     US-11-270-287-3

     Query Match             56.1%;  Score 31.8;  DB 7;  Length 516;
     Best Local Similarity   56.1%;
     Matches   60;  Conservative    2.3%;  Pred. No. 1.3;  Mismatches  47;  Indels  0;  Gaps  0;

Qy    141  CAACCTGATATATTTTAGATGCCAAATGTGCATCACCCTTTGAGTAGAATCTTGCTCAATTA  200
              ||||||||| |||||| | | |  | |||  ||  | || | || || | ||  |  ||
Db     78  CAACCTGATTTATTTTGTAAGTGCTATTGGTATAAATAATATAGGTGCAAATA  137

Qy    201  TCGGGGCCATCTTGACCGGATTAAATACACACTTATCCACCCACTGCATA  247
              | | || | ||  |||| || ||||| | |  | | |   |  | | |
Db    138  TTAGAGGGTATATAIGGATTAAATAGTTTATATTTAAATAATAGATA  184


RESULT 5
US-11-217-529-77328/c
Sequence 77328, Application US/11217529
Publication No. US20060099612A1
GENERAL INFORMATION:
;    APPLICANT: SUNTORY LIMITED
;    APPLICANT: NAKAO, YOSHIHIRO
;    APPLICANT: NAKAMURA, NORIHISA
```

```
;   APPLICANT: KODAMA, YUKIKO
;   APPLICANT: FUJIMURA, TOMOKO
;   APPLICANT: ASHIKARI, TOSHIHIKO
;   TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;   FILE REFERENCE: S-38-285
;   CURRENT APPLICATION NUMBER: US/11/217,529
;   CURRENT FILING DATE: 2005-09-02
;   PRIOR APPLICATION NUMBER: US 10/932,182
;   PRIOR FILING DATE: 2004-09-02
;   NUMBER OF SEQ ID NOS: 197023
;   SOFTWARE: Patentin version 3.3
;
;   SEQ ID NO 77328
;   LENGTH: 1632
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
US-11-217-529-77328

Query Match              2.3%;   Score 31.8;   DB 7;   Length 1632;
Matches  147;  Conservative   0;  Mismatches  192;  Indels   0;  Gaps   0;

Qy    457 GGTATCAAGCATGGCATCCATGGGAGAGATTGGGCACGTGCCTAAGGCCAAGTGCAGATAA 516
          |   ||  |     || |   |||    | |  | |  |  |  ||  |    |
Db   1212 GTTATTTATTGTGTCTTCATTGGAATAGTTAAAAACAAATTGGTTTTTCTAAATTGTATAG 1153

Qy    517 GGCATTGAGTAGGTTATGCTTGCCTTTGATTTTTAATTGATGGATTGGGTAAAAGTTTGTT 576
          | ||| |   |    | || |  |   || | || |    |  |   |  |     |  |
Db   1152 GGAACTAACCAAAAAATTCTTTCAAAATCGCTGATAATTCTATTCATGCCGCTATGTGTGAC 1093

Qy    577 TTTTTGATGATATAATGCCATACCATCACCATCAAATAAATAAAAATCTGCCGTTTG 636
          |  ||| | |    |  | |  |  | ||  || | |  | ||  | | |   |
Db   1092 TCTTGGATCCGTGTTTTTAAATGCAATAGAATTCTAAAAACTTCGTTAAAATTTTCTATTTT 1033

Qy    637 GTGGCTAAGTAAGCTATATTTTCGTTGTGTGATAAATGGGTGATTTGGGGCAAACGCTTTGTCAGTTGT 696
          |       |  |  |||  |      |   |    || | | || |  ||  |   | |
Db   1032 GGCCCTGAGTTTCCATATTCTTTGAGTTCTGAATAACTCTGCCTACGTTACTACAAAATAT 973

Qy    697 GGTCAAGCATTGCTTGGCTAAAATTGGCATCCTCTTGGTTGTAGATGATGGTGTTTGGCAATT 756
          |  | |   | ||  |  |  || | |  |  | || |     |  |  | |||| |
Db    972 ATTCAAAATAATTATTTTGTTAGGGTTGTCTCAGTTCTTGGGAGTTGAAGTGTTTTATAGTTAGTTAGC 913

Qy    757 TTCAAAAATACGCAATTTTTCGTGCCAAATAATTCCACCAT 795
          |  |  ||| | |   |   |  |   ||    |    ||
Db    912 TTCCAGAGATGGTAAAATTTCTTAAAAGATCCGTGTAACAT 874
```

US-11-217-529-3777/c

US-11-217-529-3776

```
Query Match              2.3%;   Score 31.6;   DB 7;   Length 789;
Matches  82;  Conservative   49.4%;   0;  Mismatches  84;  Indels   0;  Gaps   0;

Qy    292 ACCCAAAATCAATGCCAGTGATGATGAATGCCATAAAACCGTGCCAAGCCCATCAATTGCTGC 351
          ||   | | | ||  |  |  | | |  |   |    |  | | |  || |  |   || ||
Db    633 AGCCAAATCCTCCTCAAAATGTTGTTGTTCTTTTTCCAACTCATGAAAGGGAAAGTGATTCTCATC 574

Qy    352 AATGGTTGAGCCGATGTTTGTACCTTTTGAATCATTAAAAATATGCCTTGCTATCAATATC 411
          | |||| | || | |   | | |  | |  | |  |   | | |   | | |  |
Db    573 ATTGTTTTCATCTATGTATGTAATCATCTGAATCAGCTGAATCGGTCTGGAATGGCTTTATC 514

Qy    412 ATCAATATATTCACACGGATGCGGGTAGCCCCTTTGAAAGTCTTAAGG 457
          | | | | | | | |  ||  |   |  |  |     |  | ||
Db    513 ATTCAGTTCAATAAAGTGAATTCATTCCTTGTTATTACGGGCCATATGG 468
```

US-11-217-529-3777

US-11-217-529-3776

```
Query Match              2.3%;   Score 31.6;   DB 7;   Length 1125;
Matches  82;  Conservative   49.4%;   0;  Mismatches  84;  Indels   0;  Gaps   0;

Qy    292 ACCCAAAATCAATGCCAGTGATGATGAATGCCATAAAACCGTGCCAAGCCCATCAATTGCTGC 351
          ||   | | | ||  |  |  | | |  |   |    |  | | |  || |  |   || ||
Db    369 AGCCAAATCCTCCTCAAAATGTTGTTGTTCTTTTTCCAACTCATGAAAGGGAAAGTGATTCTCATC 310

Qy    352 AATGGTTGAGCCGATGTTTGTACCTTTTGAATCATTAAAAATATGCCTTGCTATCAATATC 411
          | |||| | || | |   | | |  | |  | |  |   | | |   | | |  |
Db    309 ATTGTTTTCATCTATGTATGTAATCATCTGAATCAGCTGAATCGGTCTGGAATGGCTTTATC 250

Qy    412 ATCAATATATTCACACGGATGCGGGTAGCCCCTTTGAAAGTCTTAAGG 457
          | | | | | | | |  ||  |   |  |  |     |  | ||
Db    249 ATTCAGTTCAATAAAGTGAATTCATTCCTTGTTATTACGGGCCATATGG 204
```

```

```
; FILE REFERENCE: EX03-003C-PC
; CURRENT APPLICATION NUMBER: US/10/504,120
; CURRENT FILING DATE: 2004-08-06
; PRIOR APPLICATION NUMBER: 60/354,824
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: 60/358,217
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,189
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,126
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/360,224
; PRIOR FILING DATE: 2002-02-26
; PRIOR APPLICATION NUMBER: 60/358,995
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/358,756
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/358,765
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/359,531
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/360,222
; PRIOR FILING DATE: 2002-02-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: Patentin version 3.2
;
; SEQ ID NO 16
; LENGTH: 3653
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-504-120-16

  Query Match   2.3%; Score 31.4; DB 6; Length 3653;
  Best Local Similarity   56.2%; Pred. No. 5.2;
  Matches   59; Conservative   0; Mismatches   46; Indels   0; Gaps   0;

Qy   753 AATTTTCAAAAATACGCAATTTTGTGCAAATAATCCACCATGCCATGCTGTGCCGATCAA 812
              | ||||||| || | ||| || | ||| || ||| || ||| | | | |
Db   2893 AAGTGTTAAAAGTACCCAATTTGAAAACCAAATGCACCCCACTACCCAGTCAGTGTGGA 2834

Qy   813 GATGATCACGGGAGAGATTTAAAAATTGTTCGCCCCTTGAGCTGCCA 857
              | ||| | ||| || ||||||||| ||| | | |||| | || ||
Db   2833 GGTGATACCAAATTGACTTTTAAAACCTTTGCTCTTCAAAGTCCA 2789
```

```
RESULT 9
US-11-217-529-166180/c
; Sequence 166180, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 166180
; LENGTH: 5026
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-166180

  Query Match    2.3%;  Score 31.4;  DB 7;  Length 5026;
  Best Local Similarity   52.7%;  Pred. No. 6.2;
  Matches   68;  Conservative   0;  Mismatches   61;  Indels   0;  Gaps   0;
```

```
Qy    530 TTATGCTTGCCTTTGATTTTTAAATTGATGATGGTAAAAGTTTGTTTTTTGATGATAT 589
          ||   |||  ||| |||    |||     |||| || |||||| ||| ||||| |||||
Db   3680 TTTTATTAGAATTAAAAATATTATTAATTACTCAGAGACTAAATAGTAGTTATTTAAAATACTAT 3621

Qy    590 AATGCCATACCATCACCATCATCAAATAAATAAATACTCAGAGACTAAATAGTAG 649
          ||  ||  || |||| || ||  |||| || || || |  |||| || ||| |||
Db   3620 AATTATTCATGATAAGGAATCATGAATTAAATATACAGATAATCCTATTTATGCTTATATTGT 3561

Qy    650 CTATTTGTT 658
          |   ||| ||
Db   3560 TGGTTTATT 3552
```

```
RESULT 10
US-11-217-529-734/c
; Sequence 734, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 734
; LENGTH: 1176
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-734

  Query Match    2.3%;  Score 31.2;  DB 7;  Length 1176;
  Best Local Similarity   50.0%;  Pred. No. 3.2;
  Matches   78;  Conservative   0;  Mismatches   78;  Indels   0;  Gaps   0;
```

```
Qy    429 GATGCGGTAGCCCTTTGAAAGTTCTTAAGGGTTATCAAGCATGGCATCCATGGGAAGATTGG 488
          | |  |  ||  ||||| ||||||  |||| |||| |  |  |||| |||||  || |||||
Db    634 GATCCAGTTTCCTTGTAAATTCCTCCAAGTTGTTAGCCCTTGACGGGCGAATTAAAATACCTGG 575

Qy    489 CAGCTGTGCCTAAGGCAAGTCCAGATAAGGCATTGAGTAGGTTATGCTTGCCCTTTGATTT 548
          ||| || |  |||| |||||  || | ||||| ||  |||| |  || | ||||| |||||
Db    574 TAATTTTTTCGTTCAGTTTCTCAGCTAAATCATCGACTTGCTTCATCATACTTCACGTC 515

Qy    549 TTAATTGATGGATTGGTTAAAAGTTTGTTTTTTGAT 584
          |||| | ||| |  |||| || |||  |||||| |||
Db    514 TTTTTCTTTTCCATTTCCAAAAGTTTCTTTTCTTTGCT 479
```

```
RESULT 11
US-11-217-529-1381/c
; Sequence 1381, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
; APPLICANT: SUNTORY LIMITED
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
```

; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
;
; SEQ ID NO 1381
; LENGTH: 1629
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-1381

Query Match   2.3%;   Score 31.2;   DB 7;   Length 1629;
Best Local Similarity 46.4%;   Pred. No. 3.9;
Matches 102;   Conservative   0;   Mismatches 118;   Indels   0;   Gaps   0;

Qy   577 TTTTTGATGATATAATAATGCCATACCATCACCATCAAAATAAATAATAAAAAATCTGCGTTTG 636
           |||| ||| | | |||||
Db   1089 TCTTGGATCGGTTTTTAAAATACAATAAAATTCCAAAAACTTCATTGAAATCTTCTATTTT   1030

Qy   637 GTGGCTAAGTAAGCTAATTATTTGTTGTCGATAATGGTGTGATTTGGGGGCAACGCTTGTCAGTGT 696
           ||| ||||
Db   1029 TGGCCCGAAGTTTCCATATCTTTCTTTGGGTCCTCGAATAGCTCTGCTACATTACTACAGATAT   970

Qy   757 TTCAAAAATACGCAAATTTTTGTGTGCAAAATAATCACCACCATG 796
           |||| |||||
Db   969 ATTCCAAAATGTTATTGGTAAGTGTTGTCAATTCTTGAGAGTTAAAGTGTGTTTATAAATTGGC   910

Qy   697 GGTCAAGCATTGCTTGGGCTAAATTGGCATCGTCTTGGTAGATGATGGCCGGTTTGGCAATT   756
           |||| ||||
Db   909 TTCCAAAAGCGATAGGTTTCTTAAGAGATCATGCAACATG 870

RESULT 12
US-11-217-529-76422/c
; Sequence 76422, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
;   APPLICANT: SUNTORY LIMITED
;   APPLICANT: NAKAO, YOSHIHIRO
;   APPLICANT: NAKAMURA, NORIHISA
;   APPLICANT: FUJIMURA, TOMOKO
;   APPLICANT: KODAMA, YUKIKO
;   APPLICANT: ASHIKARI, TOSHIHIKO
;   TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;   FILE REFERENCE: S-38-285
;   CURRENT APPLICATION NUMBER: US/11/217,529
;   CURRENT FILING DATE: 2005-09-02
;   PRIOR APPLICATION NUMBER: US 10/932,182
;   PRIOR FILING DATE: 2004-09-02
;   NUMBER OF SEQ ID NOS: 197023
;   SOFTWARE: Patentin version 3.3
; SEQ ID NO 76422
; LENGTH: 2385
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-76422

Query Match   2.3%;   Score 31.2;   DB 7;   Length 2385;
Best Local Similarity 63.2%;   Pred. No. 4.8;
Matches 48;   Conservative   0;   Mismatches 28;   Indels   0;   Gaps   0;

Qy   547 TTTTAATTGATGGATTGGTAAAAGTTGTTGTTTTTTTGATGATATAATGCCATACCATCACC 606
           |||| | |||| || | || |||||| ||||||||| | |||
Db   770 TTGAATTTGATTGTTTGGTTATTTTCCTTTGATTTCATTCGATATATTTGAC   711

Qy   607 ATCATCAAATAAAATA 622
           |||||||||
Db   710 ATAATGTGTTGAATAA 695

RESULT 13
US-11-217-529-4571/c
; Sequence 4571, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
;   APPLICANT: SUNTORY LIMITED

; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHISA
; APPLICANT: KODAMA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: S-38-285
; CURRENT APPLICATION NUMBER: US/11/217,529
; CURRENT FILING DATE: 2005-09-02
; PRIOR APPLICATION NUMBER: US 10/932,182
; PRIOR FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: Patentin version 3.3
; SEQ ID NO 4571
; LENGTH: 1077
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-4571

Query Match   2.2%;   Score 31;   DB 7;   Length 1077;
Best Local Similarity 57.9%;   Pred. No. 3.5;
Matches 55;   Conservative   0;   Mismatches 40;   Indels   0;   Gaps   0;

Qy   857 AAAATTTGAGATATGCTCAAGCTGAAAACTTGACAGCTCAAGCACCACCAAATCCATATTT 916
           ||| ||| |||| ||| |||||||| ||||| ||| |||||| ||||||||||
Db   662 AATTGTGACAACGTGCGGATCAAACTTGAAATCAAACTTGAAAAGTTTGACAGCAATTCATGTCCATATTG   603

RESULT 14
US-11-217-529-80459/c
; Sequence 80459, Application US/11217529
; Publication No. US20060099612A1
; GENERAL INFORMATION:
;   APPLICANT: SUNTORY LIMITED
;   APPLICANT: NAKAO, YOSHIHIRO
;   APPLICANT: NAKAMURA, NORIHISA
;   APPLICANT: KODAMA, YUKIKO
;   APPLICANT: FUJIMURA, TOMOKO
;   APPLICANT: ASHIKARI, TOSHIHIKO
;   TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;   FILE REFERENCE: S-38-285
;   CURRENT APPLICATION NUMBER: US/11/217,529
;   CURRENT FILING DATE: 2005-09-02
;   PRIOR APPLICATION NUMBER: US 10/932,182
;   PRIOR FILING DATE: 2004-09-02
;   NUMBER OF SEQ ID NOS: 197023
;   SOFTWARE: Patentin version 3.3
; SEQ ID NO 80459
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-11-217-529-80459

Query Match   2.2%;   Score 31;   DB 7;   Length 1593;
Best Local Similarity 57.9%;   Pred. No. 4.4;
Matches 55;   Conservative   0;   Mismatches 40;   Indels   0;   Gaps   0;

Qy   917 TGGATCGTTAATAATTCAAGTGCAGGCGTGCCAAT 951
           |||| ||| |||| |||||||||||
Db   602 TGAATCTCTATTAATGCAGCTGTTGCGGTTTGCAAT 568

RESULT 15
US-11-217-529-78136/c
; Sequence 78136, Application US/11217529

```
; Publication No. US20060099612A1
; GENERAL INFORMATION:
;   APPLICANT: SUNTORY LIMITED
;   APPLICANT: NAKAO, YOSHIHIRO
;   APPLICANT: NAKAMURA, NORIHISA
;   APPLICANT: KODAMA, YUKIKO
;   APPLICANT: FUJIMURA, TOMOKO
;   APPLICANT: ASHIKARI, TOSHIHIKO
;   TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
;   FILE REFERENCE: S-38-285
;   CURRENT APPLICATION NUMBER: US/11/217,529
;   CURRENT FILING DATE: 2005-09-02
;   PRIOR APPLICATION NUMBER: US 10/932,182
;   PRIOR FILING DATE: 2004-09-02
;   NUMBER OF SEQ ID NOS: 197023
;   SOFTWARE: PatentIn version 3.3
; SEQ ID NO 78136
;   LENGTH: 1125
;   TYPE: DNA
;   ORGANISM: Saccharomyces pastorianus
; US-11-217-529-78136

Query Match           2.2%;    Score 30.8;  DB 7;   Length 1125;
Best Local Similarity   47.0%;  Pred. No. 4.2;
Matches    95;  Conservative   0;  Mismatches 107;  Indels    0;  Gaps    0;

Qy    518 GCATTGAGTAGGTTATGCCTTGCCTTGATTTTTAATTGAATGGATTGGTAAAAGTTTGTTT 577
          ||||| ||       ||    |||   ||   ||    ||| || ||||||   |||
Db    544 GCATTTATAGCATTGCCTTCCGTTGAACATTTTTTTGAATTTTTTCGTGATACTCTGCAT 485

Qy    578 TTTTGATGATATAAATGCCATACCATCACCATCCATCAAATAAAATAAAAATCTGCGGTTTGG 637
          |   ||    |||      || |  ||    ||     ||| ||||||  |||| ||
Db    484 TTCCTTCTTTTACAATTAATCAAGTTTTCAACCAATAAACTTCAATAAATCATTCAGAGACG 425

Qy    638 TGGCTAAGTAAGCTATTTGTTGTGATAATGGTGTGATTTGGGGGCAACGCTTGTCAGTGTG 697
          |   ||||  ||||   | | ||  |||||||| ||  ||||||     ||||| || |||
Db    424 TGCTGAAGTCAATTTTAAAGTATGATAAATGCTATTCTTATTCTTTGGTATTGTTCATCTA 365

Qy    698 GTCAAGCATTGCTTGGCTAAAT 719
          |  ||   || ||   ||||||
Db    364 ATAAAAAAATCGTTCATAAAT 343
```

Search completed: May 18, 2006, 11:42:15
Job time : 52 secs